This file includes two sections. Section 1 presents comments from referees, the corresponding point-by-point responses, and the related changes in the manuscript. Section 2 is the marked-up manuscript.

Section 1: (the black font are comments from referees, the red font are authors' responses as well as the related change clarifications.)

(1) Response to comments from referee #1:

Thanks very much for your comments, suggestions and recommendation with respect to publish this paper in AMT. Our response to all your comments are listed as follows. There is an extensive discussion among the authors regarding how to revise the content. So the response is delayed, and we are sorry for this.

General comments:

The authors present a study on the influence of instrumental line shape (ILS) degradation on NDACC gas retrieval. Although this topic has been discussed in several NDACC infrared working group (IRWG) meetings in the past there is not so much in the literature, except for a few species such as ozone or water vapor. This paper describes this topic in detail for all the ten species which are mandatory to retrieve and for which a harmonized data analysis scheme is established within the IRWG. Since it is well written and gives a comprehensive presentation of the influence of an imperfect ILS I recommend publishing this paper. This paper fits in the scope of AMT and will be useful for the IRWG.

Specific comments:

- Chapter 4.3 and Table 4: Channeling error is not included in the error analysis. At least for a weak absorber such as ClONO2 this error source is not negligible.

Response: The selection of error items and their values cannot be easily standardised because most of them are instrument/site dependent. In this paper, we already included most common error items in the error analysis. The channeling error was not included because: 1, it is instrument dependent and it is not a common error, some instrument may have very weak channeling effect; 2, The main point of this paper is the same regardless of including or not including channeling error. This is because error analysis is the post processing (last step) of NDACC retrieval, how many errors

to be included may have influence on the total error and thus the fractional difference of statistic errors, but have no influence on the total column, DOFs and profile which are obtained before post processing step.

Related change: None

- While Haidinger fringes are presented for scenarios in Figs. 11 & 13 Haidinger fringes are missing for those in Fig. 1.

Response: We have included Haidinger fringes for those in Fig.1.

Related change: In the revised paper, Fig.2 showing Haidinger fringes for those in Fig.1 is included.

- The conclusion (as well as in the abstract) 'For total column retrieval, the stratospheric gases are more sensitive to instrumental line shape degradation than the tropospheric gases.' is a bit qualitative. I would suggest to add some numbers: For typical misalignment scenarios the column of O3, HCl, HF and ClONO2 changed by 3, 6, 5 and 35%, respectively.

Response: We have improved this description as your suggestion.

Related change: A quantitative description "For a typical ILS degradation (10%), the total columns of stratospheric gases O_3, HNO_3, HCl, HF, and $ClONO_2$ changed by 1.9%, 0.7%, 4%, 3%, and 23%, respectively. While the columns of tropospheric gases CH_4, CO, N_2O, C_2H_6, and HCN changed by 0.04%, 2.1%, 0.2%, 1.1%, and 0.75%, respectively." have been included in both abstract and conclusion. Please check the conclusion and abstract sections for details.

- Table 5 nicely summarizes the recommendations for ME. I would suggest to add a sentence to the end of the abstract and the conclusion summarizing this result: 'For the retrieval of NDACC standard stratospheric species a ME within +-5% is required. Therefore, the alignment of an NDACC instrument needs to be better than 5% in terms of ME' or something similar.

Response: We have added some sentences to summarize Table 4 (i.e., Table 5 in the previous version).

Related change: Some sentences to summarize Table 4 (Table 5 in the previous version) are added, please check section 6 for details.

Technical corrections:

- p. 5, line 139: increasing misalignment with increasing opd

Response: We have revised this sentence as your suggestion.

Related change: Now it is "Typically, the increasing misalignment with increasing OPD (b, f, h or i) causes negative ME amplitude and the decreasing misalignment with increasing OPD (e, g or j) causes positive ME amplitude."

- Legend of Figs. 5&6 and x axis description in Fig. 9 are hard to read (at least in my hardcopy).

Response: In order to make the content more concise on the main point of this paper and catch attention to the influence of ILS on the retrieval of gases, following referee #2' suggestion, we have removed Figs. 5&6 which contribute trivial to the main point of this paper. We have updated Fig.9, now x axis description is very clear.

Related change: We removed Figs. 5&6 and updated the x axis of Fig.9.

- Figs.: 'l' in small letter in HCl and ClONO2

Response: All "HCL" and "$CLONO_2$" in this paper have been changed to "HCl" and "$ClONO_2$", respectively.

Related change: We have revised all "HCL" and "$CLONO_2$" as "HCl" and "$ClONO_2$", respectively.

(2) Response to comments of referee #2:

Thanks very much for your comments, suggestions and recommendation. Our response to all your comments are listed as follows. There is an extensive discussion among the authors regarding how to revise the content. So the response is delayed, and we are sorry for this.

This paper "The influence of instrumental line shape degradation on NDACC gas retrievals" by Sun et al presents sensitivity studies regarding the influence of ILS degradation in FTIR retrievals results. It is well known that the shape of the gas absorption lines can be impacted by the ILS, if instruments are not well-aligned. However, up to now there are not many published results regarding the quantitative impact. The lack of details in the impact of ILS makes this study important. The topic

of the study is interesting and suitable for the journal. I suggest some revisions before its publication.

General Comments:

- The manuscript is short and lack important quantitative details regarding the finding with respect to the influence of ILS (results section). While the reader can check figures, and make sense of quantitative results authors do not explain in detail in the text their findings (see specific comments below).

Response: In the revised paper, we have added quantitative details regarding the finding of this study.

Related change: We have added quantitative details regarding the finding of this study in abstract, sections 6 and 7.

- Authors use an ideal ILS of actual FTIR measurements to know the influence of different ILS degradation. This statement is important since all quantitative Figures shown are with respect to this reference. However, there is a lack of proof about the ILS of actual measurements. If the ILS of actual measurements deviate from ideal the results shown here might change significantly. I suggest to include the actual ILS of the FTIR and its temporal variability.

Response: The Hefei site has run NDACC observations with the Bruker 125HR for more than three years. We regularly use a low-pressure HBr cell to diagnose the misalignment of the spectrometer and to realign the instrument when indicated. As shown in Fig.5 (new added), all actual ILS degradations of the FTIR spectrometer within this selected period are less than 2% and can be regarded as ideal. The paper focuses on relative % difference of each quantity, the influence due to this assumption is negligible.

On the other hand, if the ILS of actual measurements deviate from ideal, it will cause an offset to both X and X_{ref} in equation 9. Assuming the offset is Δ, equation 9 becomes,

$$D = \frac{(X \pm \Delta)-(X_{ref} \pm \Delta)}{(X_{ref} \pm \Delta)} *100 = \frac{X - X_{ref}}{(X_{ref} \pm \Delta)} *100$$

Considering that X_{ref} is close to ($X_{ref} \pm \Delta$), the results (especially the trends) shown here won't change significantly.

Related change: We have included the actual ILS of the FTIR and its temporal variability, i.e., Fig.5 in the revised version.

- (a) My understanding based on the analysis and table 5 is that the effect of the ILS (given that degradation of ILS for most FTIRs-NDACC is low) can be regarded as negligible for most gases, except, N_2O, correct?. (b) I do not find suggestions beside table 5 for including the ILS in the analysis of standard NDACC gases. (c) Given that the ILS effect is negligible, would you suggest using ideal ILS?. (d) I recommend to include a section with specific recommendations for FTIR/NDACC sites that will bring dialogue towards and harmonization in ILS.

Response:

(a) If total column is the target, the findings are the effect of the ILS degradation **cannot** be regarded as negligible for most gases, except, N_2O and CH_4. Your understanding is up-side-down.

(b) In the revised paper, we have added quantitative details regarding the finding of this study. Please check abstract, sections 6 and 7 for details.

(c) If total column is the target, the ILS effect is not negligible for most gases, we suggest to keep the ILS degradation of each site within the recommendation. Note that the retrievals of certain gases, e.g., O_3, CH_4, CO, and N_2O, can be divided into multiple independent sub layers depending on total DOFs. The recommendation don't apply to partial column integrated over each sub layer because, as Figs. 17 and 18 show, the sensitivity of profile to ILS degradation is altitude dependent. How ILS degradation influences partial column of each NDACC gas and how much ILS deviation from unity is acceptable if an ideal line shape is assumed beyond the scope of this paper and will be published elsewhere. Details can be found in section 6.

(d) In the revised paper, we have added quantitative details regarding the finding of this study. Please check abstract, sections 6 and 7 for details.

Related change: We have added quantitative details regarding the finding of this study in abstract, sections 6 and 7.

- I recommend a thorough revision in format/style of the citations.

Response: This has been done.

Related change: We have updated the format/style of the citations.

- I recommend a thorough English revision.

Response: The revised version is already gone through a copy-editing service.

Related change: A copy-editing service has been used.

Specific Comments:

Abstract:

L27, I would change "current NDACC gases" with "current standard NDACC gases" since FTIR retrievals go beyond these mandatory gases.

Response: This has been done in the revised version.

Related change: "current NDACC gases" becomes "current standard NDACC gases"

L33-34, influence is written twice, remove one.

Response: This has been done in the revised version.

Related change: We removed one of them

L38-40, "In order to suppress the influence on total column for ClONO2 and other NDACC gases within 10% and 1%, respectively, the permitted maximum ILS degradation for each NDACC gas was deduced (summarized in Table 5)". In my opinion, authors should summarize table 5 in the abstract rather than pointing the reader to table 5. I found this difficult to interpret if the reader aims to check the abstract only.

Response: In the revised paper, we have added quantitative details regarding the finding of this study.

Related change: We have added quantitative details regarding the finding of this study in abstract, sections 6 and 7.

Introduction:

L61, "FTIR spectrometers are highly precise and stable measurement devices and the instrumental line shapes (ILSs) not far from the theoretical limit if carefully aligned". This sentence is not clear. Please change it accordingly. Consider something like this:

“FTIR spectrometers are highly precise and stable devices and if carefully aligned the instrumental line shape (ILS) might not be far from the theoretical limit”.

Response: We have revised this sentence as your suggestion.

Related change: We have revised this sentence as your suggestion.

L72-74. It might be important to mention that TCCON only uses NIR, fewer gases, and only columns are aimed compared to NDACC.

Response: This has been done as your suggestion.

Related change: A new sentence “The TCCON network only operates in near infrared (NIR) region and aims at column of fewer gases. While the NDACC network operates in both NIR and mid-infrared (MIR) regions and aims at both columns and profile of many gases.” has been inserted in this section.

3 Simulation of ILS degradation

I could not find a description of ALIGN60 in the references provided. I suggest to describe in more detail ALIGN60 in this paper.

Response: A more detailed descriptions of ALIGN60 has been included.

Related change: A more detailed descriptions of ALIGN60 provided by the ALIGN60 developer has been included in this paper, please check section 3 in the revised version for details.

NDACC gases retrieval

4.1 Retrieval strategy

- L151-152. Is there a reference for the retrieval setting of NDACC?, if so cite it here.

Response: The retrieval setting of NDACC can be found via the link “(https://www2.acom.ucar.edu/irwg/links)”.

Related change: We have cited the above link in the revised version.

-The size of Table 2 can be significantly smaller. I suggest to remove all cells that are similar for the different gases and add a description in either the main text or caption of table, e.g., spectroscopy, P,T profiles, etc

Response: All parameters that are the same for different gases are removed from Table 2, but the descriptions in the main text are kept. Now the size of Table 2 is significantly smaller than previous version.

Related change: We significantly shorten Table 2.

4.2 Averaging kernels

- There are 26 Figures in the main text and I would consider removing some, e.g., Fig. 3 and 4 provide similar information. I would remove Fig 3 (or move it to supplemental information).

Response: The previous Fig.3 is removed.

Related change: Previous Fig.3 is removed in the revised version.

- Change to appropriate chemical formulas, e.g., HCL to HCl, etc.

Response: This has been done.

Related change: All "HCL" and "$CLONO_2$" in this paper have been changed to "HCl" and "$ClONO_2$", respectively.

4.3 Error Analysis

I would expect a description of the ILS in the uncertainty budget here. However, it is not clear how the error in ILS influences the uncertainty budget in either table 2 or Figs 5 and 6.

Response: Detailed descriptions regarding how the error in ILS influences the quantities such as the total column, RMS, random uncertainty, systematic uncertainty, total uncertainty, DOFs, and profile as well as how much is acceptable can be found in the discussion, i.e., section 6. They are the purpose of this paper, should be present after the investigation (sensitivity study) rather than present before the investigation.

Related change: None

- (a) In order to catch attention to the influence of ILS in the retrieval of gases I would remove Figs. 5 and 6. (b) Again, I think 26 Figs are overwhelming. (c) Instead, in table 4, which also does not add information, add quantitative numbers of leading/dominant errors including the ILS uncertainty.

Response:

(a) Both figures have been removed in the revised version.

(b) Previous Figs.3, 5, and 6 have been removed, Figs.21- 26 tell the similar information and have been replaced by only one figure. However, as the referees' suggestions, we included one figure showing the actual ILS degradation and one

figure showing Haidinger fringes of Fig.1. Now they are in total 20 figures.

(c) In order to catch attention to the influence of ILS in the retrieval of gases, we have significantly shorten the auxiliary contents. The error analysis may be important for study that focuses on retrieval itself, but contributes trivially to the main point of this paper. Table 4 has been removed in the revised version.

Related change: Figs.3, 5, and 6 have been removed, Figs.21- 26 have been replaced by one figure. One new figure showing the actual ILS degradation and one new figure showing Haidinger fringes of Fig.1 are added. Auxiliary contents, e.g., error analysis part have been shorten significantly and Table 4 has been removed.

5. ILS influence

- It is mention that the ILS degradation of the FTIR at Hefei is less than 2% but authors do not show how they infer this. This is key in order to avoid convolution problems with the different types of degradation.

Response: The Hefei site has run NDACC observations with the Bruker 125HR for more than three years. We regularly use a low-pressure HBr cell to diagnose the misalignment of the spectrometer and to realign the instrument when indicated. As shown in Fig.5, all actual ILS degradations of the FTIR spectrometer within this selected period are less than 2% and can be regarded as ideal. The paper focuses on relative % difference of each quantity, the influence due to this assumption is negligible.

Related change: We have included the actual ILS of the FTIR and its temporal variability, i.e., Fig.5 in the revised version.

- L243-247. It is not clear whether a single spectrum is used (what time, sza, conditions?) or all spectra recorded on Feb 16, 2016. Clarify.

Response: Unlike TCCON network, the NDACC has seven consecutive optical filters to reduce the broadband signal (avoiding detector non-linearity) and it is not possible to retrieve all ten mandatory species within one filter spectra. For your comments, our answers are: the statistical analysis of each gas is based on a single spectrum, but different gas may use different filter spectrum. Thus, 5 spectra in total are used here but all of them are recorded on Feb. 16, 2016.

Related change: A statement has been include in section 5.

- L255-265. Expand a description of the different filter criteria used here. It is clear that retrievals need to converge, but what about the 3% rms limit, what does SIV mean and why 10% is used?

Response: These criteria are used to remove those spectra that have sampling errors or contaminated by aerosols, clouds, hazes or other unpredictable objects which cause a low SNR or a large detecting intensity variation. These spectra normally show bad fitting RMSs or in accuracy retrievals. We have included this clarification in the revised paper.

Related change: An expand description has been included here.

- The color code of ME amplitude, PE, etc in Figs. 7 and 8 are different. To be consistent, change to same color code scheme. Remove the ideal case in Fig. 8.

Response: These problems have been solved and now the two figures are consistent.

Related change: We updated the two figures, please check Figs.5 and 6 in the revised version for details.

- Do results shown in Figs 7 and 8 correspond to a single spectrum? if so include date/time in the caption.

Response: The NDACC has seven consecutive optical filters to reduce the broadband signal (avoiding detector non-linearity) and it is not possible to retrieve all ten mandatory species within one filter spectra. The statistical analysis of each gas is based on a single spectrum, but different gas may use different spectrum. Thus, 5 spectra in total are used here but all of them are recorded on Feb. 16, 2016.

Related change: A statement has been include in section 5.

- It is quite strange that % difference in total columns (Fig. 7) is larger than % Difference of profiles in Fig. 8. Maybe Fig.8 is only the fraction?

Response: After a careful check to our python scripts, we found we forgot to include a factor of 100 (see equation (9)) in calculations of fractional difference in profile, but all other quantities don't have this problem. Yes, all profile related figures in previous version do indicate the fraction rather than % difference. However, in the revised version, a factor of 100 has been included and the problems in all profile related

figures have been solved. Thanks very much for pointing out these problems.

Related change: All profile related figures have been multiplied by a factor of 100.

- Use appropriate name for gases, e.g., change HCL to HCl, etc.

Response: This has been done.

Related change: All "HCL" and "$CLONO_2$" in this paper have been changed to "HCl" and "$ClONO_2$", respectively.

5.1 ME amplitude and PE influence

- It is interesting to see in Fig. 7 that for some gases the % difference in RMS is negative, which would mean that the RMS of the reference is greater than using ILS degradation. Why would the rms be smaller using degraded ILS if the FTIR is characterized as ideal?

Response: This "abnormal" phenomenon also troubled us for quite a long time during proceeding this study. Finally, we were lucky to figure out the reasons. This is because, for certain cases, the number of iterative step (to get converge) is different when using different ILS. For an example, the CO retrieval may converge with 3 iterative steps with the ideal ILS, but may need 4 or 5 iterative steps to get converge with degraded ILS. Thus, it is possible that the RMS of the reference is greater than using ILS degradation. This phenomenon do not occur quite often but indeed exist.

Related change: None

- In general, there is a lack of description in findings here. I recommend to have a more quantitative analysis and description of results in this section

Response: In the revised paper, we have added quantitative details regarding the finding of this study.

Related change: We have added quantitative details regarding the finding of this study in abstract, sections 6 and 7.

Section 2: marked up file, as follows

In brief: Figs.3, 5, and 6 have been removed, Figs.21- 26 have been replaced by one figure. One new figure showing the actual ILS degradation and one new figure showing Haidinger fringes of Fig.1 are added. Auxiliary contents, e.g., error analysis

part have been shorten significantly and Table 4 has been removed. We have added quantitative details regarding the finding of this study in abstract, sections 6 and 7. Detailed description of ALIGN60 is added. Other minor revisions responded to referees' comments also performed

The marked up file is as follow, please check the red underlined sentences for details :

The influence of instrumental line shape degradation on NDACC gas retrievals: total column and profile

Youwen Sun[1, 3)+], Mathias Palm [2)+], Cheng Liu [3, 4, 1)1], Frank Hase [5)], David Griffth [6)], Christine Weinzierl [2)], Christof Petri [2)], Wei Wang [1)], and Justus Notholt [2)]

(1 *Key Laboratory of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei 230031, China*)

(2 *University of Bremen, Institute of Environmental Physics, P. O. Box 330440, 28334 Bremen, Germany*)

(3 *Center for Excellence in Urban Atmospheric Environment, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China*)

(4 *University of Science and Technology of China, Hefei, 230026, China*)

(5 *Karlsruhe Institute of Technology (KIT), Institute for Meteorology and Climate Research (IMK-ASF), Karlsruhe, Germany*)

(6 *School of Chemistry, University of Wollongong, Northfields Ave, Wollongong, NSW, 2522, Australia*)

+These two authors contributed equally to this work

Abstract:

We simulated Instrumental line shape (ILS) degradations with respect to typical types of misalignment, and compared their influence on each NDACC (Network for Detection of Atmospheric Composition Change) gas. The sensitivities of total column, root mean square of fitting residual (RMS), total random uncertainty, total systematic uncertainty, total uncertainty, degrees of freedom for signal (DOFs), and profile with respect to different levels of ILS degradation for all current standard NDACC gases, i.e., O_3, HNO_3, HCl, HF, $ClONO_2$, CH_4, CO, N_2O, C_2H_6, and HCN, were investigated. The influence of an imperfect ILS on NDACC gases retrieval were assessed, and the

Correspondence to: Cheng Liu (chliu81@ustc.edu.cn)

consistency under different meteorological conditions and solar zenith angles (SZA) were examined. The study concluded that the influence of ILS degradation can be approximated by the linear sum of individual modulation efficiency (ME) amplitude influence and phase error (PE) influence. The PE influence is of secondary importance compared with the ME amplitude. Generally, the stratospheric gases are more sensitive to ILS degradation than the tropospheric gases, and the positive ME influence is larger than the negative ME. For a typical ILS degradation (10%), the total columns of stratospheric gases O_3, HNO_3, HCl, HF, and $ClONO_2$ changed by 1.9%, 0.7%, 4%, 3%, and 23%, respectively. While the columns of tropospheric gases CH_4, CO, N_2O, C_2H_6, and HCN changed by 0.04%, 2.1%, 0.2%, 1.1%, and 0.75%, respectively. In order to suppress the fractional difference in total column for $ClONO_2$ and other NDACC gases within 10% and 1%, respectively, the maximum positive ME degradations for O_3, HNO_3, HCl, HF, $ClONO_2$, CO, C_2H_6, and HCN should be less than 6%, 15%, 5%, 5%, 5%, 5%, 9%, and 13%, respectively; the maximum negative ME degradations for O_3, HCl, and HF should be less than 6%, 12%, and 12%, respectively; the influence of ILS degradation on CH_4 and N_2O can be regarded as negligible.

Key words: NDACC, FTIR, Instrumental line shape, Profile retrieval

1 Introduction

In order to achieve consistent results between different FTIR (Fourier transform infrared) sites, the TCCON (Total Carbon Column Observing Network, http://www.tccon.caltech.edu/) and NDACC (Network for Detection of Atmospheric Composition Change, http://www.ndacc.org/) have developed strict data acquisition and retrieval methods to minimize site to site differences (Hase et al., 2012; Wunch et al., 2010 and 2011; Washenfelder, 2006; Messerschmidt et al., 2010; Kurylo, 1991; Davis et al., 2001; Schneider, et al.,2008; Kohlhepp et al., 2011; Hannigan et al., 2009; Vigouroux et al., 2008 and 2015). Interferograms are acquired with similar instruments operated with common detectors, acquisition electronics and/or optical

filters. These interferograms are first converted to spectra and then these spectra are analyzed using dedicated processing algorithms, i.e., GFIT, PROFFIT or SFIT (Wunch et al., 2010 and 2015; Hase et al., 2006; Hannigan and Coffey, 2009). Typically, the TCCON network only uses the Bruker 125HR instruments (http://www.tccon.caltech.edu/; https://www.bruker.com/) with specified settings (entrance aperture, amplification of the detected signal). In the NDACC network, other instruments are used as well, e.g., the Bruker M series, a BOMEM DA8 in Toronto, Canada and a self-built spectrometer in Pasadena, USA (http://www.ndacc.org/; https://www.bruker.com/). FTIR spectrometers are highly precise and stable devices, and if carefully aligned, the instrumental line shape (ILS) might not be far from the theoretical limit. However, their alignment can change abruptly as a consequence of operator intervention or drift slowly due to mechanical degradation over time (Olsen et al., 2004; Duchatelet et al., 2010; Hase et al., 2012; Feist et al., 2016). Moreover, the NDACC observation may change the entrance field stop size if incident radiation changes. This practice may introduce a dependency of the instrument alignment status on the optical settings because the mechanical errors between different field stops may be non-negligible and inconsistent (Sun et al., 2017). Biases between sites would arise if all these misalignments are not properly characterized.

The TCCON network only operates in near infrared (NIR) region and aims at column of fewer gases. While the NDACC network operates in both NIR and mid-infrared (MIR) regions and aims at both columns and profile of many gases. The TCCON assumes an ideal ILS in spectra retrieval, and the maximum ILS degradation is prescribed as 5% for the modulation efficiency (ME) amplitude (Wunch et al., 2011 and 2015). This assumption still holds within the required accuracy of the results. In the NDACC gases retrieval, the ILS can be assumed as ideal if spectrometer is well aligned, or if misalignment exists, described by LINEFIT results derived from dedicated cell measurements or retrieved together with the gas profile from an atmospheric spectrum using a polynomial (Vigouroux et al., 2008 and Vigouroux et al., 2015). How these ILS treatments influence the NDACC gases retrieval and how

much ILS deviation from unity is acceptable for each NDACC gas if an ideal line shape is assumed are still not fully quantified, and it may be better to assume an ideal ILS. The practice of co-retrieving ILS parameters from atmospheric spectra without dedicated cell measurements is not to be recommended because the observed shapes of spectral lines are exploited primarily for inferring the vertical distribution of the trace gases, the ILS and the trace gas profiles have similar effects on the line shape, i.e., changing the shape and width of the line. Overlapping lines, i.e., due to interfering gases may introduce an asymmetry in the absorption lines which may be undistinguishable from an ILS phase deviation.

This paper investigates the influence of ILS degradation on total column and profile of current standard NDACC gas retrievals and deduces the maximum ILS deviations allowable for suppressing the influence within a specified acceptable ranges.

2 Characteristics of ideal and imperfect ILSs

The ILS is the Fourier transform of the weighting applied to the interferogram. This weighting consists of two parts: an artificially applied part to change the calculated spectrum and an unavoidable part which is due to the fact that the interferogram is finite in length (box car function), the divergence of the beam is non-zero (due to the non-zero entrance aperture), and several other effects which are due to misalignment (Davis et al., 2001, chapter 9). The ILS consisting of only the unavoidable parts of the line shape is called the ideal line shape.

The theoretical ideal ILS as defined in equation (3), when the instrument is well aligned, is a convolution of sinc and rectangular functions (defined in equations (1) and (2)), representing the finite length of the interferogram and the finite circular field of view (FOV) of the spectrometer (Davis et al., 2001).

$$SINC(\sigma, L) = 2L \frac{\sin(2\pi\sigma L)}{2\pi\sigma L} \tag{1}$$

$$RECT(\sigma, \sigma_0, \theta) = \begin{cases} \dfrac{2}{\sigma_0 \theta^2} & if \ -0.5\sigma_0\theta^2 \le \sigma \le 0 \\ 0 & otherwise \end{cases} \quad (2)$$

$$ILS(\sigma, \sigma_0, L, \theta) = SINC(\sigma, L) * RECT(\sigma, \sigma_0, \theta) \quad (3)$$

where σ is the wavenumber, σ_0 is the central wavenumber, L is the optical path difference (OPD) and θ is the angular radius of the circular internal FOV of the spectrometer. For standard NDACC measuring conditions, $L \geq 180$ cm and θ defined by the entrance field stop size in the light path.

The LINEFIT software calculates the deviation of the measured ILS from the ideal ILS (Hase et al., 2001 and 2012). It retrieves a complex ME as a function of OPD, which is represented by a ME amplitude and a phase error (PE) (Hase et al., 1999). The ME amplitude is connected to the width of the ILS while the PE quantifies the degree of ILS asymmetry. For a perfectly aligned spectrometer, it would meet the ideal nominal ILS characteristics if smear and vignetting effects were neglected, and thus have an ME amplitude of unity and a PE of zero along the whole interferogram. However, if a FTIR spectrometer is subject to misalignment, the ME amplitude would deviate from unity and the PE deviate from zero (Hase et al., 2012). This results in an imperfect ILS.

3 Simulation of ILS degradation

We use the program ALIGN60 to simulate ILS degradation in a high resolution FTIR spectrometer typically used in the NDACC network. As an auxiliary tool of LINEFIT, ALIGN60 is a raytracing model for FTIR spectrometers following the classical Michelson design, assuming one fixed and one movable arm, and using cube corners instead of plane mirrors. It calculates the resulting phase distortions in the recombined beam and from these deduces the variable intensity observed by the detector. ALIGN60 takes into account the lateral shear error of the movable retro-reflector as function of OPD, a decenter of the field stop with respect to the optical axis, an unsharp boundary line or deformation of the field stop image (as possibly caused by a defocused collimator), and vignetting effects with increasing

OPD. It can generate trustworthy results with respect to all types of misalignment (Hase et al., 1999). In this simulation, the entrance beam section was assumed to be circular with a diameter of 8.0 cm. The ILS was only calculated from positive side of interferogram. The smear and vignetting effects were not taken into account. The misalignment of a FTIR spectrometer can be expressed via two perpendicular axes perpendicular to the beam direction. For a circular entrance beam, the same misalignment in either direction results in a similar ILS. Thus, this work only considers misalignment in one axis.

The misalignments as inputs of ALIGN60 are listed in Table 1, the resulting ILSs are shown in Fig. 1, and the corresponding Haidinger fringes at the maximum OPD are shown in Fig. 2. The ME deviation, decenter of Haidinger fringes and ILS deterioration varying over misalignment are evident. All types of misalignment cause nonlinear ME deviations except decentering of measuring laser (*c*) and the constant shear (*d*) which mainly affect PE and result in linear PE deviation. Two types of ILS degradation are evident, one is referred to as positive ME and has a ME amplitude of larger than unity. The other one is referred to as negative ME and has a ME amplitude of less than unity. Typically, the increasing misalignment with increasing OPD (*b*, *f*, *h* or *i*) causes negative ME amplitude and the decreasing misalignment with increasing OPD (*e*, *g* or *j*) causes positive ME amplitude. For the same misalignment amplitude, the decreasing misalignment causes more ME deviation than the increasing misalignment. Regardless of positive or negative ME, the ME deviation shape depends on misalignment type and the same misalignment amplitude causes the same deviation in ME amplitude. The decentering of the entrance filed stop is equivalent to the linear increasing misalignment.

4 NDACC gases retrieval

4.1 Retrieval strategy

The influence of ILS degradation on all current standard NDACC gases, i.e., O_3, HNO_3, HCl, HF, $ClONO_2$, CH_4, CO, N_2O, C_2H_6, and HCN, is investigated. Typical atmospheric vertical profiles of these gases are shown in Fig.3. There are five

stratospheric gases and five tropospheric gases. The retrieval settings for all these gases as recommended by the NDACC are listed in Table 2 (https://www2.acom.ucar.edu/irwg/links). The latest version of profile retrieval algorithm SFIT4 v 0.9.4.4 is used (http://www.ndacc.org/). The basic principle of SFIT4 is using an optimal estimation technique for fitting calculated-to-observed spectra (Rodgers, 2000; Hannigan and Coffey, 2009). All spectroscopic line parameters are adopted from HITRAN 2008 (Rothman et al., 2009). This might not be ideal, but we keep it to achieve consistent results. A priori profiles of pressure, temperature and water vapor for the measurement days are interpolated from the National Centers for Environmental Protection and National Center for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al., 1996). A priori profiles of the target gases and the interfering gases except H_2O use the WACCM4 (Whole Atmosphere Community Climate Model) model data. We follow the NDACC standard convention with respect to micro windows (MWs) selection and the interfering gases consideration (https://www2.acom.ucar.edu/irwg/links). For the interfering molecules that affect the target gas retrieval, H_2O should be treated with care as it is almost always present in all MWs, to varying degrees. It has been dealt with differently for different gas. For HNO_3 and $ClONO_2$, H_2O is treated as the other interfering species: only a scaling of a single *a priori* profile is made. For other gases, the H_2O profile is retrieved simultaneously with the target gas profile. No de-weighting signal to noise ratios (SNR) are used except for CO and HCl which utilize a de-weighting SNR of 500 and 300, respectively.

The selection of the regularization (a priori covariance matrix $\mathbf{S}_a$ and SNR) cannot be easily standardised because it depends on the real variability for each gas. In optimal estimation, the selection of $\mathbf{S}_a$ is very important in the inversion process and, together with the measurement noise error covariance matrix $\mathbf{S}_\varepsilon$, will lead to the following averaging kernel matrix **A** (Rodgers, 2000):

$$\mathbf{A} = \mathbf{G}_y\mathbf{K}_x = (\mathbf{K}_x^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}_x^T + \mathbf{S}_a^{-1})^{-1}\mathbf{K}_x^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}_x \quad (4)$$

where $\mathbf{G}_y$ is the sensitivity of the retrieval to the measurement. $\mathbf{K}_x$ is weighting function matrix or Jacobian matrix that links the measurement vector $\boldsymbol{y}$ to the state vector $\boldsymbol{x}$: $\Delta\boldsymbol{y}=\mathbf{K}_x\Delta\boldsymbol{x}$. **A** characterizes the vertical information contained in the FTIR retrievals. In this study, we assume $\mathbf{S}_\varepsilon$ to be diagonal and its diagonal elements are the inverse square of the SNR. The vertical information content of the retrieved target gas

profile can be quantified by the number of degrees of freedom for signal (DOFs), which is the trace of **A**, defined in Rodgers (2000) by:

$$d_s = tr(\mathbf{A}) = tr((\mathbf{K}_x^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}_x^T + \mathbf{S}_a^{-1})^{-1} \mathbf{K}_x^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}_x^T) \quad (5)$$

The diagonal elements of $\mathbf{S}_a$ represent the assumed variability of the target gas volume mixing ratio (VMR) at a given altitude, and the off diagonal elements represent the correlation between the VMR at different altitudes. We can see in Table 3 that, except CO and HCN, the target gases are using an a priori covariance matrix with diagonal elements constant with altitude corresponding to 10, 20, 50 or 100 % variability; the largest variability are for HNO_3, HCl and $ClONO_2$. For CO, the diagonal elements of $\mathbf{S}_a$ correspond to 27% from ground to 34 km and decrease down to 11% at the top of atmosphere. For HCN, the diagonal elements of $\mathbf{S}_a$ correspond to 79% from ground to 5 km and decrease down to 21% at the top of atmosphere. No correlation of off diagonal matrix elements is used in all retrievals except for $ClONO_2$ which uses exponential correlation with a HWHM (half with at half-maximum) of 8 km. The SNR values for all retrievals are the real values taken from each individual spectrum. The ILSs for all retrievals are using the simulations in section 3.

4.2 Averaging kernels

The rows of **A** are the so called averaging kernels and they represent the sensitivity of the retrieved profile to the real profile. Their FWHM is a measure of the vertical resolution of the retrieval at a given altitude. The area of averaging kernels represents sensitivity of the retrievals to the measurement. This sensitivity at altitude *k* is calculated as the sum of the elements of the corresponding averaging kernels, $\sum_i \mathbf{A}_{ki}$. It indicates the fraction of the retrieval at each altitude that comes from the measurement rather than from the a priori information (Rodgers, 2000). A value close to zero at a certain altitude indicates that the retrieved profile at that altitude is nearly independent of measurement and is therefore approaching the a priori profile.

The averaging kernels and their areas for these ten NDACC gases are shown in Fig. 4. The altitude ranges with sensitivity larger than 0.5 and the corresponding total DOFs are summarized in Table 3. These sensitive ranges indicate that the retrieved profile information comes by more than 50% from measurement, or, in other words, that the a priori information influences the retrieval by less than 50%. Each gas has different sensitive range. The sensitive range for HCN, CO and C_2H_6 is mainly

tropospheric, and for $ClONO_2$, HCl and HF is mainly stratospheric. O_3, CH_4 and N_2O have high retrieval sensitivity in both troposphere and stratosphere. The HNO_3 has high retrieval sensitivity in stratosphere and in atmospheric boundary layer below 1.5 km.

4.3 Error analysis

As listed in Table 2, we classified errors as systematic or random according to whether they are constant between consecutive measurements, or vary randomly. For comparison, the error items considered in error analysis are the same for the retrieval of all gases. The smoothing error $\mathbf{E}_s$ is calculated via equation (6), the measurement error $\mathbf{E}_m$ is calculated via equation (7), and all other error items $\mathbf{E}_{var}$ are calculated via equation (8) (Rodgers, 2000).

$$\mathbf{E}_s = (\mathbf{A} - \mathbf{I})\mathbf{S}_a(\mathbf{A} - \mathbf{I})^T \tag{6}$$

$$\mathbf{E}_m = \mathbf{G}_y\mathbf{S}_\varepsilon\mathbf{G}_y^T \tag{7}$$

$$\mathbf{E}_{var} = \mathbf{G}_y\mathbf{K}_{var}\mathbf{S}_{var}\mathbf{K}_{var}^T\mathbf{G}_y^T \tag{8}$$

where $\mathbf{S}_{var}$ is the error covariance matrix of *var*. $\mathbf{K}_{var}$ is weighting function matrix of *var*. Here *var* refers to one of the error items in Table 2 except smoothing error and measurement error. In this study, the *a priori* error covariance for all non-retrieval parameters are set the same for all gases retrieval.

5 ILS influence study

This section presents the ILS influence study, whereby the degraded ILSs that simulated by ALIGN60 are used in the SFIT forward model, and the fractional difference (D%) in various quantities for each gas relative to the retrieval with an ideal ILS are computed. For each gas, sections 5.1 and 5.2 only select one typical spectrum for study. In order to retrieve these ten gases, five spectra with different wavenumber coverage are used. All of them are randomly selected from the routine measurements on a clear day at Hefei on February 16, 2016. The consistency of the resulting deduction is evaluated in section 5.3 where one year of measurements from August 2015 to August 2016 are used. The Hefei site has run NDACC observations with the Bruker 125HR for more than three years. We regularly use a low-pressure HBr cell to diagnose the misalignment of the spectrometer and to realign the

instrument when indicated. As shown in Fig.5, all actual ILS degradations of the FTIR spectrometer within this selected period are less than 2% and can be regarded as ideal. For all spectra used in this study, the retrievals with all levels of ILS degradation fulfill the following filter criteria:

1) The root mean squares (RMSs) of the residual (difference between measured and calculated spectra after the fit) in all fitting windows has to be less than 3%.

2) The retrievals should converge for all levels of ILS degradation.

3) The concentrations of the target and interfering gases at each sub layer should be positive.

4) The solar intensity variation (SIV) should be less than 10%. The SIV within the duration of a spectrum is the ratio of the standard deviation to the average of the measured solar intensities.

These criteria are used to remove those spectra that have sampling errors or contaminated by aerosols, clouds, hazes or other unpredictable objects which cause a low SNR or a large detecting intensity variation. In following calculations, we have taken the retrievals with an ideal ILS as the reference. The fractional difference is defined here as,

$$\mathrm{D}\% = \frac{\mathrm{X} - \mathrm{X}_{ref}}{\mathrm{X}_{ref}} \times 100 \quad (9)$$

where **X** is a vector which can include multiple elements such as gas profile or only one element such as DOFs, RMS, total column, total random uncertainty, total systematic uncertainty, or total uncertainty. The total random uncertainty and systematic uncertainty are the sum in quadrature of each individual uncertainty listed in Table 2, and the total uncertainty is the sum in quadrature of total random uncertainty and total systematic uncertainty. $\mathbf{X}_{ref}$ is the same as **X** but for the nominal ideal ILS.

5.1 ME amplitude and PE influence

In order to determine how the ILS degradation affects the NDACC gas retrievals, the results deduced from ILS considering both ME amplitude and PE are compared to

those only considering ME amplitude or PE. All types of ILS degradation in section 3 are used in this study. Fig.6 exemplifies the case of ILS *j*, where the differences in total column, RMS, random uncertainty, systematic uncertainty, total uncertainty, and DOFs for each gas relative to the retrieval with an ideal ILS are compared. Fig.7 shows the fractional difference in profile of each gas for ILS *j*. The results show that the influence of ILS degradation on the total column, RMS, random uncertainty, systematic uncertainty, total uncertainty, DOFs, and profile can be approximated by the linear sum of individual ME amplitude influence and PE influence. The PE influence is of secondary importance compared with the ME amplitude influence. The comparisons for the results retrieved with ILS *a* to *i* come to the same conclusions.

Figs.8 and 9 show the influence of ILS *a* to *j* on total column and profile of all NDACC gases. The resulting influence amounts depend on deviation amount and deviation shape of ME. For positive MEs, in most cases, the ILS *j* causes the maximum influence, and for negative MEs, the ILS *i* causes the maximum influence. In a real instrument, the misalignment is a combination of misalignment *a* to *j*. In principle, for the same misalignment amplitude, it should not cause influence exceeding misalignment *i* or *j*. In the following, misalignment *i* and *j* are selected on behalf of negative and positive ME respectively to investigate how the ILS degradation influence the NDACC gas retrievals.

5.2 Sensitivity study

We simulated seven levels of negative ME *i* and positive ME *j* with ALIGN60, and incorporated them in the SFIT forward model, and then calculated the fractional difference in various quantities for each gas relative to the retrieval with an ideal ILS. The misalignments as inputs of ALIGN60 and the resulting ILSs are shown in Figs. 10 and 12. The corresponding Haidinger fringes at the maximum misalignment position are shown in Figs. 11 and 13. The ME deviation, decenter of Haidinger fringes and ILS deterioration varying over misalignment are evident. Fig.14 is the sensitivity of total column with respect to different levels of ILS degradation. Figs. 15 ~ 18 are the same as Fig. 14 but for DOFs, RMS, uncertainty and profile. The results

show that the ILS degradation affected total column, RMS, DOFs, retrieval uncertainty, and profile. Generally, the larger the ME deviation, the larger the influence. The positive and negative ME have opposite influence on total column, DOFs, total uncertainty and profile.

With respect to total column, the influence of ILS degradation on stratospheric gases is generally larger than the tropospheric gases. For a typical ILS degradation (10%), the total columns of stratospheric gases O_3, HNO_3, HCl, HF, and $ClONO_2$ changed by 1.9%, 0.7%, 4%, 3%, and 23%, respectively. While the total columns of tropospheric gases CH_4, CO, N_2O, C_2H_6, and HCN changed by 0.04%, 2.1%, 0.2%, 1.1%, and 0.75%, respectively. For O_3 and HNO_3, positive ME causes an overestimated total column and negative ME causes an underestimated total column. For other gases, negative ME causes an overestimated total column and positive ME causes an underestimated total column. For all gases except O_3 and CH_4, the positive ME influence is larger than the negative ME influence. For CH_4, the negative ME influence is larger than the positive ME influence. For O_3, the level of the positive ME influence and the negative ME influence is very close.

For all gases, positive ME increases the DOFs and negative ME decreases DOFs. For all gases except HF and CH_4, both positive ME and negative ME increase RMS. For HF, positive ME increases RMS while negative ME decreases RMS. For CH_4, positive ME decreases RMS and negative ME increases RMS.

The influence on systematic uncertainty and random uncertainty depends on ME deviation type and gas type. The influence on total uncertainty is the combination of the influence on total systematic uncertainty and total random uncertainty. For all gases except O_3, positive ME decreases total uncertainty and negative ME increases total uncertainty. For O_3, positive ME increases total uncertainty and negative ME decreases total uncertainty.

The ILS degradation causes an evident difference in profile within the altitude ranges that show high retrieval sensitivity in Fig.4, or in other words, the sensitive ranges listed in Table 3. Generally, the profile is more sensitive to positive ME than negative PE, and the influence of ILS degradation on stratospheric gases is larger than

the tropospheric gases.

5.3 consistency evaluation

This section uses the spectra recorded at Hefei from August 2015 to August 2016 to evaluate the consistency of above study. These spectra span a large difference in atmospheric water vapor, SZAs, surface pressures, surface temperatures, wind speeds, and wind directions (Fig. 19). All retrievals fulfill the above filter criteria are included in this study. A simulated ILS *j* with maximum ME amplitude deviation of 5% is used in the retrieval. The results are compared to the retrievals deduced from an ideal ILS.

Fig. 20 exemplifies the fractional difference in total column, RMS, total uncertainty, and DOFs for each gas as a function of SZA. The results show that the fractional difference in total column, total uncertainty, and DOFs for all gases are consistent under different SZAs. For most gases, the fractional difference in RMS exhibits more scatters than the total column, total uncertainty, and DOFs. However, they are independent of SZA, and most of them are less than 10%. In general, the influence of ILS degradation on NDACC gases retrieval shows good consistency under different SZAs. The fractional difference as functions of humidity, pressure, SZA, temperature, wind direction, and wind speed come to the same conclusions.

6 Discussion and recommendation

For each gas, the *a priori* covariance matrices of $\mathbf{S}_a$, $\mathbf{S}_\varepsilon$, and $\mathbf{S}_{var}$ are the same in the aforementioned study. According to equations 6 ~ 8, we conclude that the ILS degradation altered the weighting function matrix $\mathbf{K}_x$ and eventually altered the quantities such as the total column, RMS, random uncertainty, systematic uncertainty, total uncertainty, DOFs, and profile. The change of $\mathbf{K}_x$ is attributed to the fact that the ILS degradation alters gas absorption line shape and hence alters the structure of calculated spectra, and aggravates the mismatch between the calculated spectra and the measured spectra.

The stratospheric gases are more sensitive to ILS degradation than the tropospheric gases, and the $ClONO_2$ exhibits the largest sensitivity. This is because the absorption structure in stratosphere is narrower than that in troposphere, and is

more easily affected by ILS degradation. We set the acceptable fractional difference in total column for $ClONO_2$ and other NDACC gases as 10% and 1%, respectively. Considering an excessively large of ME degradation (e.g., > 20%) seldom occurred within NDACC network because of the regular alignment at each site, the permitted maximum ILS degradation for each gas is deduced in Table 5 as:

1) The influence of ILS degradation on CH_4 and N_2O can be regarded as negligible.

2) If a misalignment causes positive ME degradation, the maximum degradations for O_3, HNO_3, HCl, HF, $ClONO_2$, CO, C_2H_6, and HCN should be less than 6%, 15%, 5%, 5%, 5%, 5%, 9%, and 13%, respectively.

3) If a misalignment causes negative ME degradation, the maximum degradations for O_3, HCl, and HF should be less than 6%, 12%, and 12%, respectively.

Note that the retrievals of certain gases, e.g., O_3, CH_4, CO, and N_2O, can be divided into multiple independent sub layers depending on total DOFs. The above deductions don't apply to partial column integrated over each sub layer because, as Figs. 17 and 18 show, the sensitivity of profile to ILS degradation is altitude dependent. How ILS degradation influences partial column of each NDACC gas and how much ILS deviation from unity is acceptable if an ideal line shape is assumed beyond the scope of this paper and will be published elsewhere.

7 Conclusion

We assessed the influence of instrumental line shape degradation on all current NDACC gases retrieval via investigation of sensitivities of total column, root mean square of fitting residual, total random uncertainty, total systematic uncertainty, total uncertainty, degrees of freedom, and profile with respect to modulation efficiency degradations. The study concluded that the influence of instrumental line shape degradation can be approximated by the linear sum of individual modulation efficiency amplitude influence and phase error influence. The phase error influence is of secondary importance compared with the modulation efficiency amplitude influence. The influence amounts depend on deviation amount and deviation shape of

the modulation efficiency.

The stratospheric gases are more sensitive to instrumental line shape degradation than the tropospheric gases, and the positive modulation efficiency has more influence on total column or profile than the negative modulation efficiency. For a typical ILS degradation (10%), the columns of stratospheric gases O_3, HNO_3, HCl, HF, and $ClONO_2$ changed by 1.9%, 0.7%, 4%, 3%, and 23%, respectively. While the columns of tropospheric gases CH_4, CO, N_2O, C_2H_6, and HCN changed by 0.04%, 2.1%, 0.2%, 1.1%, and 0.75%, respectively. The influence of instrumental line shape degradation on NDACC gas retrievals shows good consistency under different meteorological conditions and solar zenith angle. In order to suppress the fractional difference in total column for $ClONO_2$ and other NDACC gases within 10% and 1%, respectively, the maximum positive modulation efficiency degradations for O_3, HNO_3, HCl, HF, $ClONO_2$, CO, C_2H_6, and HCN should be less than 6%, 15%, 5%, 5%, 5%, 5%, 9%, and 13%, respectively; the maximum negative modulation efficiency degradations for O_3, HCl, and HF should be less than 6%, 12%, and 12%, respectively; the influence of ILS degradation on CH_4 and N_2O can be regarded as negligible.

7 Acknowledgements

This work is jointly supported by the National High Technology Research and Development Program of China (No. 2016YFC0200800, 2016YFC0203302, 2016YFC0200404, No. 2017YFC0210002), the National Science Foundation of China (No. 41605018, No.41775025, No. 41405134, No. 41575021, No. 51778596, No. 91544212, No. 41722501), Anhui Province Natural Science Foundation of China (No. 1608085MD79), and the German Federal Ministry of Education and Research (BMBF) (Grant No. 01LG1214A). The processing environment of SFIT4 and some plot programs are provided by National Center for Atmospheric Research (NCAR), Boulder, Colorado, USA. The NDACC networks are acknowledged for supplying the SFIT software and advice.

8 References

Davis, S. P., Abrams, M. C., and Brault, J. W.: Fourier transform spectrometry,

Academic Press, ISBN: 0-12-042510-6, 2001.

Duchatelet P., Demoulin P., Hase F., Ruhnke R., Feng W., Chipperfield M. P., Bernath P. F., Boone C. D., Walker K. A., and Mahieu E.: Hydrogen fluoride total and partial column time series above the Jungfraujoch from long term FTIR measurements: Impact of the line shape model, characterization of the error budget and seasonal cycle, and comparison with satellite and model data, J. Geophys. Res., 115, D22306, doi:10.1029/2010JD014677, 2010.

Feist D. G., Arnold S. G., Hase F., and Ponge D.: Rugged optical mirrors for Fourier transform spectrometers operated in harsh environments, Atmos. Meas. Tech., 9, 2381–2391, www.atmos-meas-tech.net/9/2381/2016/ doi:10.5194/amt -9-2381-2016, 2016.

Hannigan, J. and Coffey, M.: semiautonomous FTS observation system for remote sensing of stratospheric and tropospheric gases, Journal of Atmospheric and Oceanic Technology . 09/2009; 26(9). DOI: 10.1175/2009JTECHA1230.1

Hase, F., Demoulin, P., Sauval, A., Toon, G. C., Bernath, P., Goldman, A., Hannigan, J., Rinsland C.: An empirical line-by-line model for the infrared solar transmittance spectrum from 700 to 5000 cm^{-1}, J. Quant. Spectrosc. Radiat. Transfer, 2006, 102, 450 - 463.

Hase, F.: Improved instrumental line shape monitoring for the ground-based, high-resolution FTIR spectrometers of the Network for the Detection of Atmospheric Composition Change, Atmos. Meas. Tech., 5, 603–610, doi:10.5194/amt-5-603-2012,2012.

Hase, F., Drouin, B. J., Roehl, C. M., Toon, G. C., Wennberg, P. O., Wunch, D., Blumenstock, T., Desmet, F., Feist, D. G., Heikkinen, P., De Mazière, M., Rettinger, M., Robinson, J., Schneider, M., Sherlock, V., Sussmann, R., Té, Y., Warneke, T., and Weinzierl, C.: Calibration of sealed HCl cells used for TCCON instrumental line shape monitoring, Atmos. Meas. Tech., 6, 3527-3537, doi:10.5194/amt-6-3527-2013, 2013.

Hase, F., Blumenstock, T., and Paton-Walsh, C.: Analysis of the instrumental line shape of high-resolution Fourier transform IR spectrometers with gas cell

measurements and new retrieval software, Appl. Optics, 38, 3417–3422, 1999.

Kalnay E., Kanamitsu M., Kistler R., et al. (1996) The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society, 77, 437-472.

Kurylo, M. J.: Network for the detection of stratospheric change (NDSC), SPIE Proceedings 1991, P. Soc. Photo-Opt. Ins., 1491, 168–174, 1991.

Kohlhepp, R., Barthlott, S., Blumenstock, T., Hase, H., Kaiser, I., Raffalski, U., and Ruhnke, R.: Trends of HCl, $ClONO_2$, and HF column abundances from ground-based FTIR measurements in Kiruna (Sweden) in comparison with KASIMA model calculations, Atmos. Chem. Phys., 11, 4669–4677, 2011 www.atmos-chem-phys.net/11/4669/2011/doi:10.5194/acp-11-4669-2011

Messerschmidt, J., Macatangay, R., Notholt, J., Petri, C., Warneke, T., and Weinzierl, C.: Side by side measurements of CO_2 by ground-based Fourier transform spectrometry (FTS), Tellus B, 62, 749–758, oi:10.1111/j.1600-0889.2010.00491.x ,2010.

Olsen, S. C. and Randerson, J. T.: Differences between surface and column atmospheric CO_2 and implications for carbon cycle research, J. Geophys. Res.-Atmos., 109, D02301, doi:10.1029/2003JD003968, 2004.

Schneider, M., Redondas, A., Hase, F., Guirado, C., Blumenstock, T., and Cuevas, E.: Comparison of ground-based Brewer and FTIR total column O_3 monitoring techniques, Atmos. Chem. Phys., 8, 5535–5550, doi:10.5194/acp-8-5535-2008, 2008.

Rodgers, C. D.: Inverse methods for atmospheric sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics, Vol. 2, World Scientific Publishing Co., Singapore, 2000.

Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J.-P., Chance, K., Coudert, L. H., Danaj, V., Devi, V. M., Fally, S., Flaud, J.-M., Gamache, R. R., Goldmanm, A., Jacquemart, D., Kleiner, I., Lacome, N., Lafferty, W. J., Mandin, J.-Y., Massie, S. T., Mikhailenko, S. N., Miller, C. E., Moazzen-Ahmadi, N., Naumenko, O. V., Nikitin, A. V., Orphal, J., Perevalov, V. I., Perrin, A., Predoi-Cross, A., Rinsland, C.

P., Rotger, M., Šime. cková, M., Smith, M. A. H., Sung, K., Tashkun, S. A., Tennyson, J., Toth, R. A., Vandaele, A. C., and Vander Auwera, J.: The Hitran 2008 molecular spectroscopic database, J. Quant. Spectrosc. Ra., 110, 533–572, 2009.

Vigouroux, C., De Mazière, M., Demoulin, P., Servais, C., Hase, F.,Blumenstock, T., Kramer, I., Schneider, M., Mellqvist, J., Strandberg, A., Velazco, V., Notholt, J., Sussmann, R., Stremme, W., Rockmann, A., Gardiner, T., Coleman, M., and Woods, P.: Evaluation of tropospheric and stratospheric ozone trends over Western Europe from ground-based FTIR network observations, Atmos. Chem. Phys., 8, 6865–6886, 2008, http://www.atmos-chem-phys.net/8/6865/2008/.

Vigouroux, C., Blumenstock, T., Coffey, M., Errera, Q., Garcá, O., Jones, N. B., Hannigan, J. W., Hase, F., Liley, B., Mahieu, E., Mellqvist, J., Notholt, J., Palm, M., Persson, G., Schneider, M., Servais, C., Smale, D., Thölix, L., and De Mazière, M.: Trends of ozone total columns and vertical distribution from FTIR observations at eight NDACC stations around the globe, Atmos. Chem. Phys., 15, 2915-2933, doi:10.5194/acp-15-2915-2015, 2015.

Washenfelder, R. A.: Column abundances of carbon dioxide and methane retrieved from ground-based near-infrared solar spectra, PhD thesis, California Institute of Technology, Pasadena, California (available at: http://thesis.library.caltech.edu), 2006.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmos. Meas. Tech., 3, 1351–1362, doi:10.5194/amt-3-1351-2010, 2010.

Wunch, D., Toon, G. C., Blavier, J.-F. L.,Washenfelder, R., Notholt, J., Connor, B. J.,

Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The Total Carbon Column Observing Network, Phil. T. Roy. Soc. A, 369, 2087–2112, doi:10.1098 /rsta.2010.0240, 2011.

Wunch, D., Toon G. C., Sherlock V., Deutscher N. M., Liu C., Feist D. G., and Wennberg P. O.: The Total Carbon Column Observing Network's GGG2014 Data Version. 10.14291/tccon.ggg2014.documentation.R0/1221662, 2015.

Sun, Y., Palm, M., Weinzierl, C., Petri, C., Notholt, J., Wang, Y., and Liu, C.: Technical note: Sensitivity of instrumental line shape monitoring for the ground-based high-resolution FTIR spectrometer with respect to different optical attenuators, Atmos. Meas. Tech., 10, 989-997, doi: 10.5194/amt-10-989-2017, 2017.

9 Figs

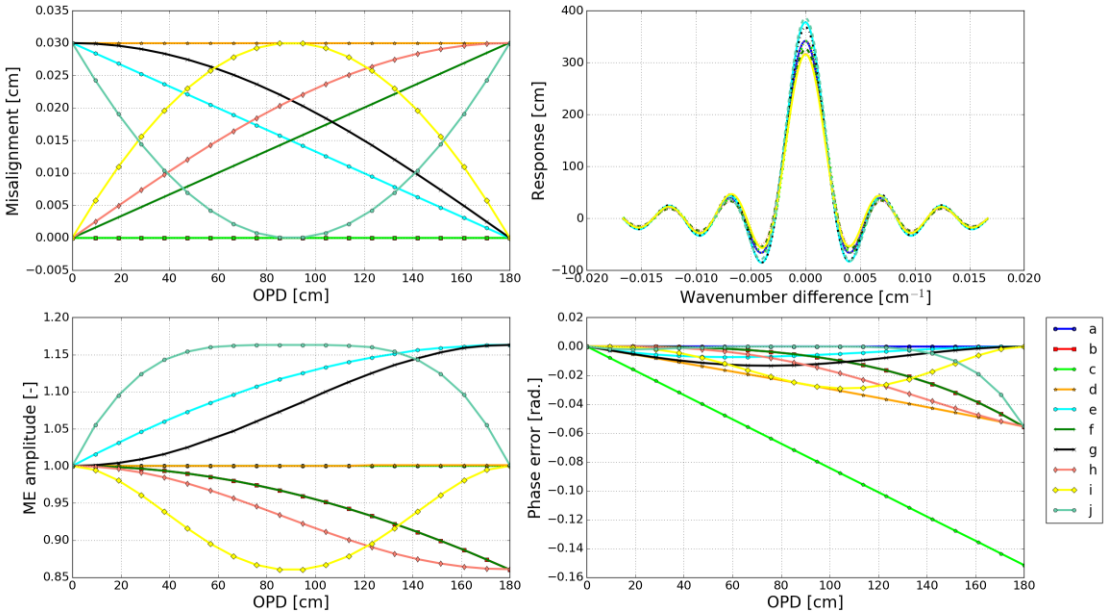

Fig.1. Simulated ILS degradation with respect to different types of misalignment. The results are derived from ALIGN60. Top left demonstrates different types of misalignment (*a* to *j*) used in the simulation, top right is the resulting ILS, bottom left is the resulting ME amplitude, and bottom right is the resulting PE. Descriptions for the misalignment *a* to *j* are listed in Table 1.

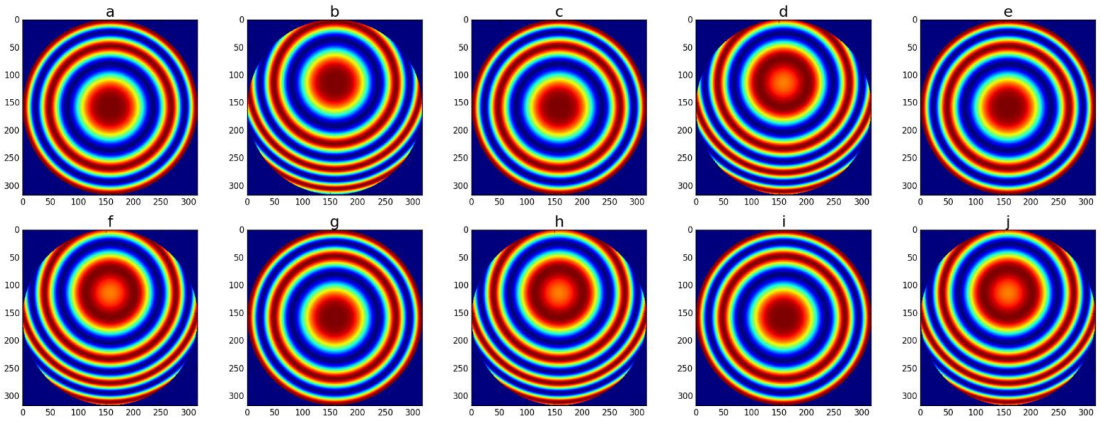

Fig.2. The Haidinger fringes at maximum OPD for misalignment *a* to *j* shown in Fig. 1.

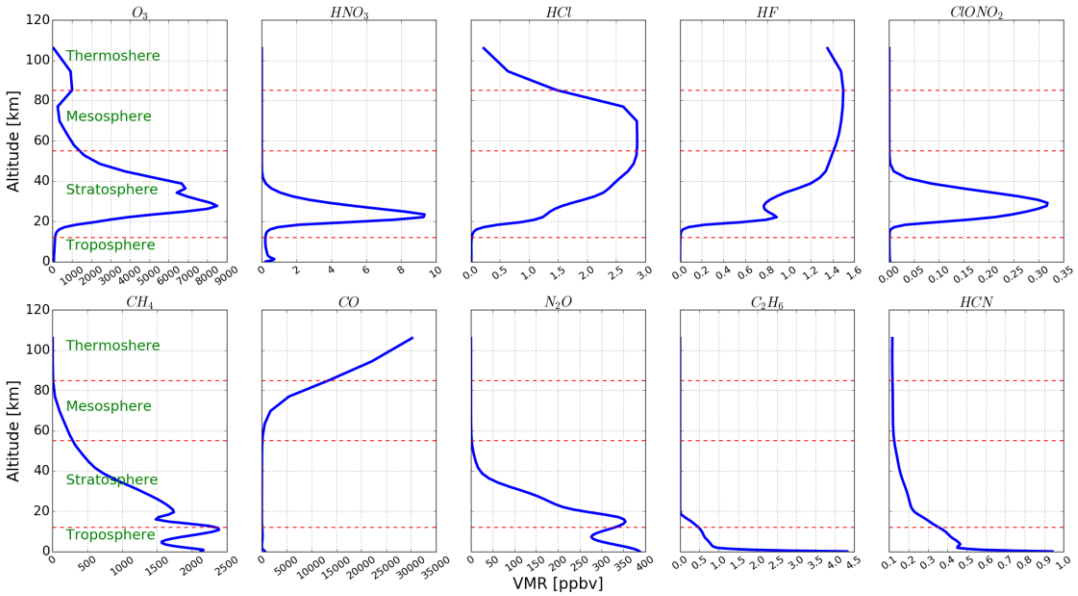

Fig.3. Typical profiles of ten NDACC gases. Bottom panels are five tropospheric gases, i.e., CH_4, CO, N_2O, C_2H_6, and HCN. Top panels are five stratospheric gases, i.e., O_3, HNO_3, HCl, HF, and $ClONO_2$. Although the CO concentration above 60 km is much higher than that in the troposphere, it is regarded as tropospheric gas because it is an anthropologic pollution gas and shows large variation in troposphere.

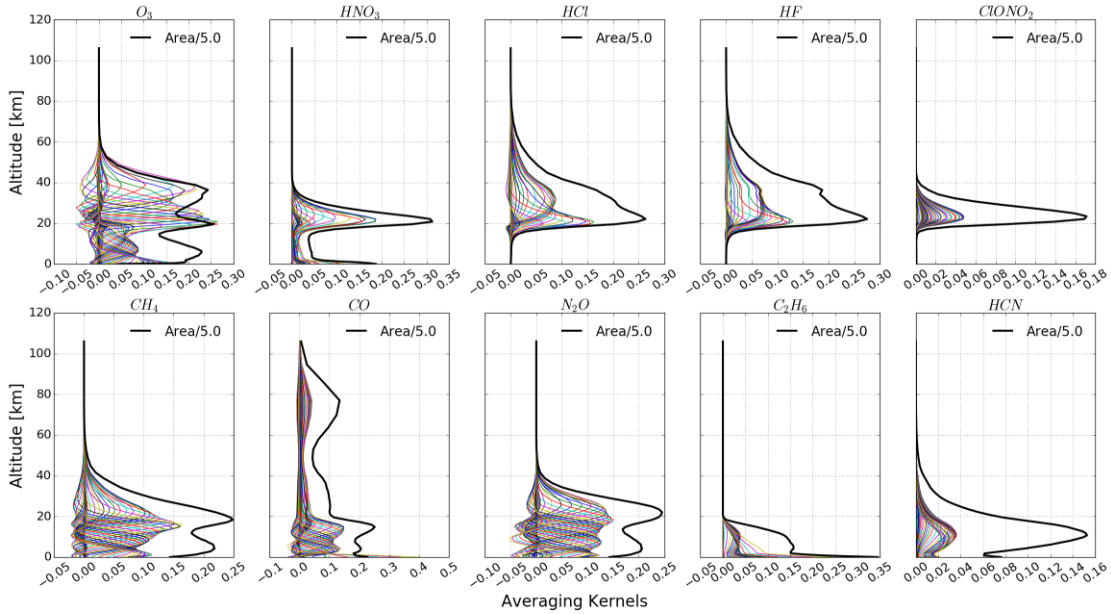

Fig.4. Averaging kernels of ten NDACC gases (color fine lines), and their area scaled by a factor of 0.2 (black bold line). They are deduced from the spectra recorded at Hefei on February 16, 2016 with an ideal ILS.

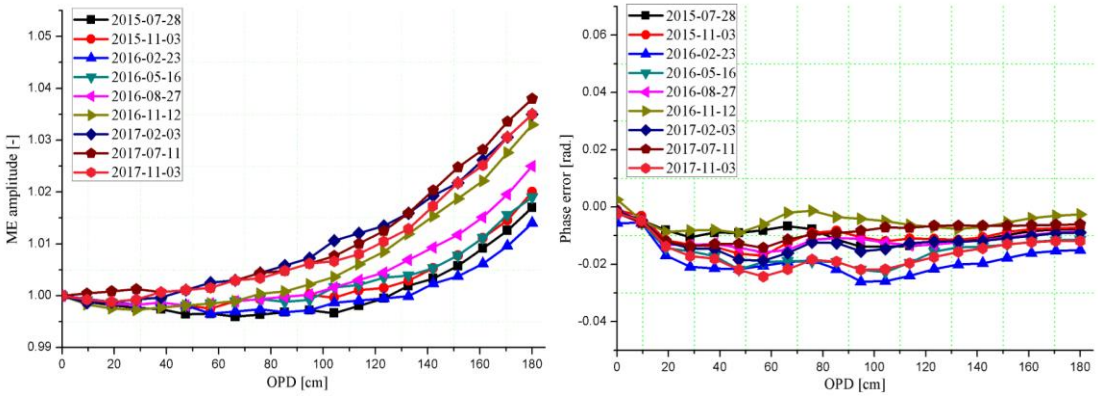

Fig.5. ME amplitudes (left) and phase errors (right) along with OPD deduced from HBr cell measurements at Hefei.

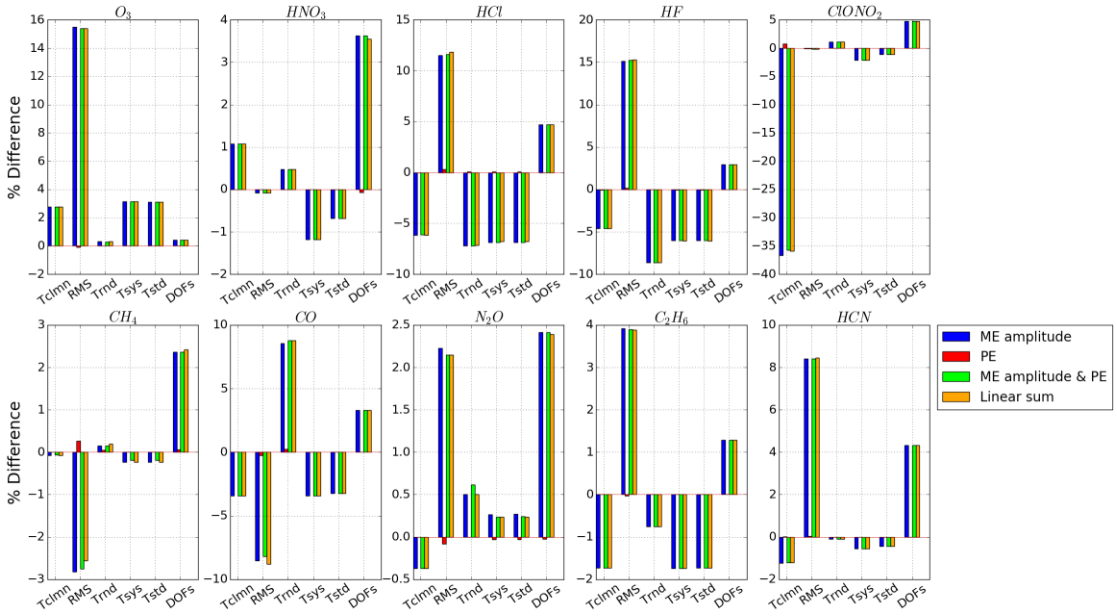

Fig.6. Fractional difference in total column, RMS, total random uncertainty, total systematic uncertainty, total uncertainty, and DOFs for misalignment *j*. "ME amplitude" represents the ILS only taken ME amplitude deviation into account. "PE" represents the ILS only taken PE deviation into account. "ME amplitude & PE" represents the ILS taken both ME amplitude and PE deviations into account. "Linear sum" represents the fractional difference of each item is linear sum of "ME amplitude" and "PE". The ME amplitude and PE are obtained from ALIGN60 with misalignment *j* in Fig.1. The results are deduced from the spectra recorded at Hefei on February 16, 2016.

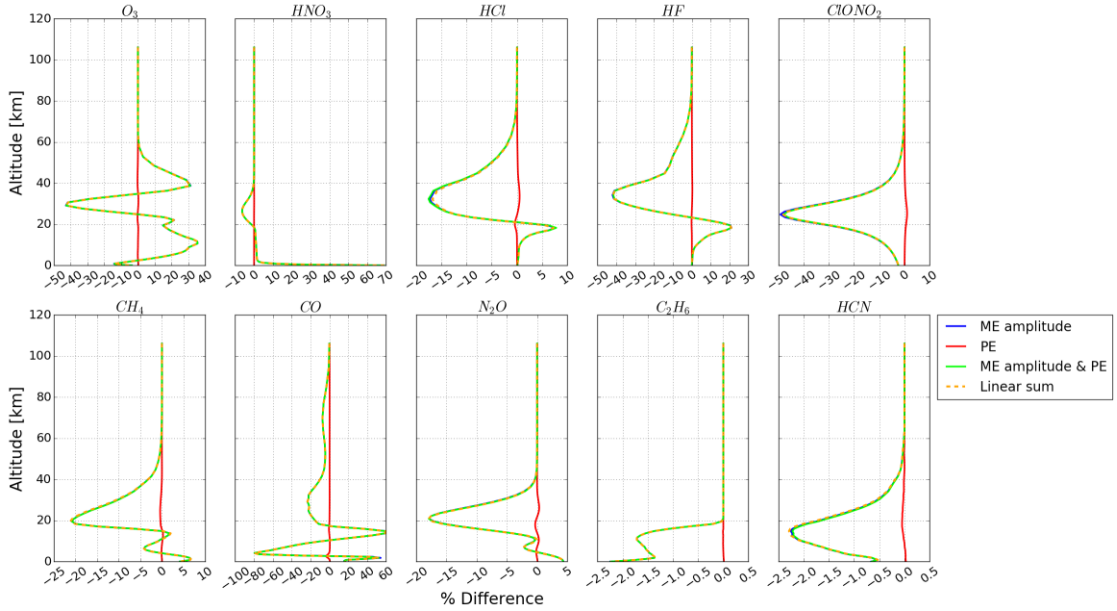

Fig.7. Fractional difference in profile for misalignment *j*. The nomenclatures in the plot legend is same as Fig.6. The results are deduced from the spectra recorded at Hefei on February 16, 2016.

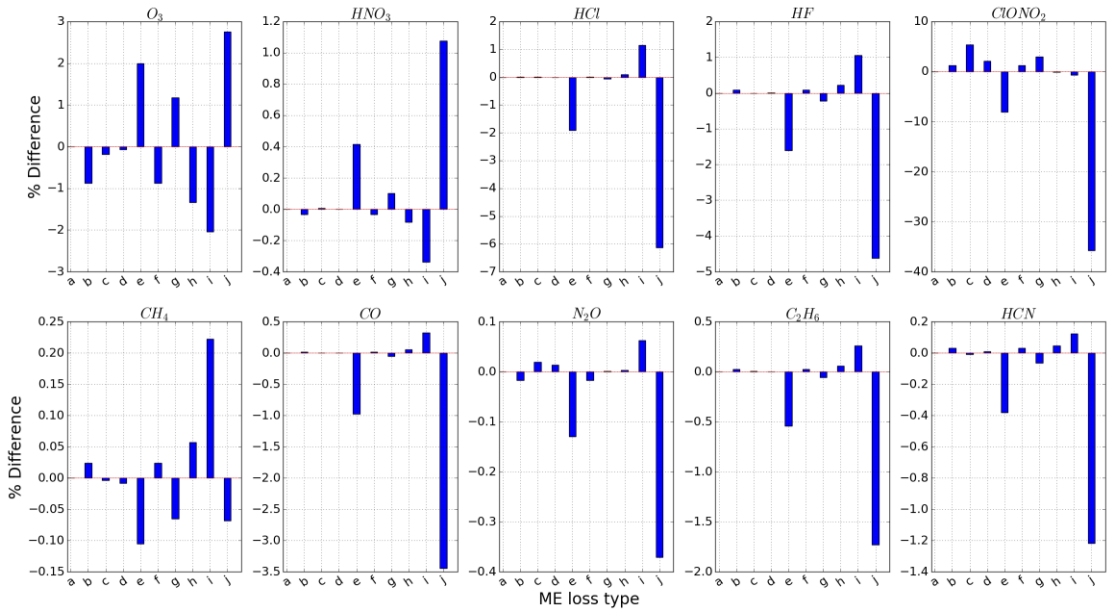

Fig.8. Sensitivity of total column to different types of ILS degradation. The ILS *a* to *j* correspond to misalignment *a* to *j* in Table1. The results are deduced from the spectra recorded at Hefei on February 16, 2016.

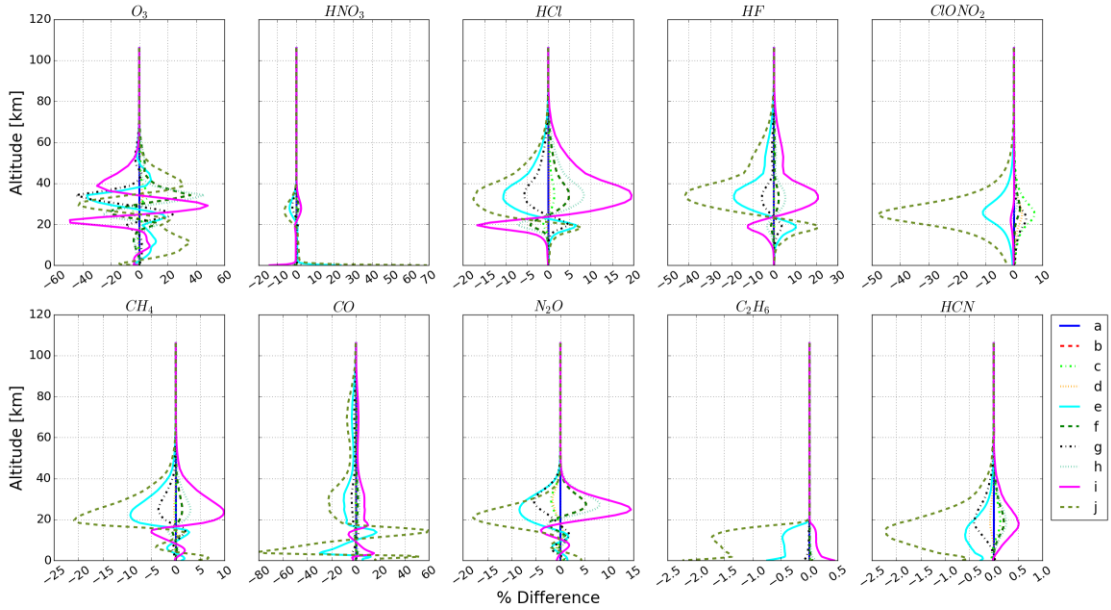

Fig.9. Sensitivity of profile to different types of ILS degradation. The ILS *a* to *j* correspond to misalignment *a* to *j* in Table1. The results are deduced from the spectra recorded at Hefei on February 16, 2016.

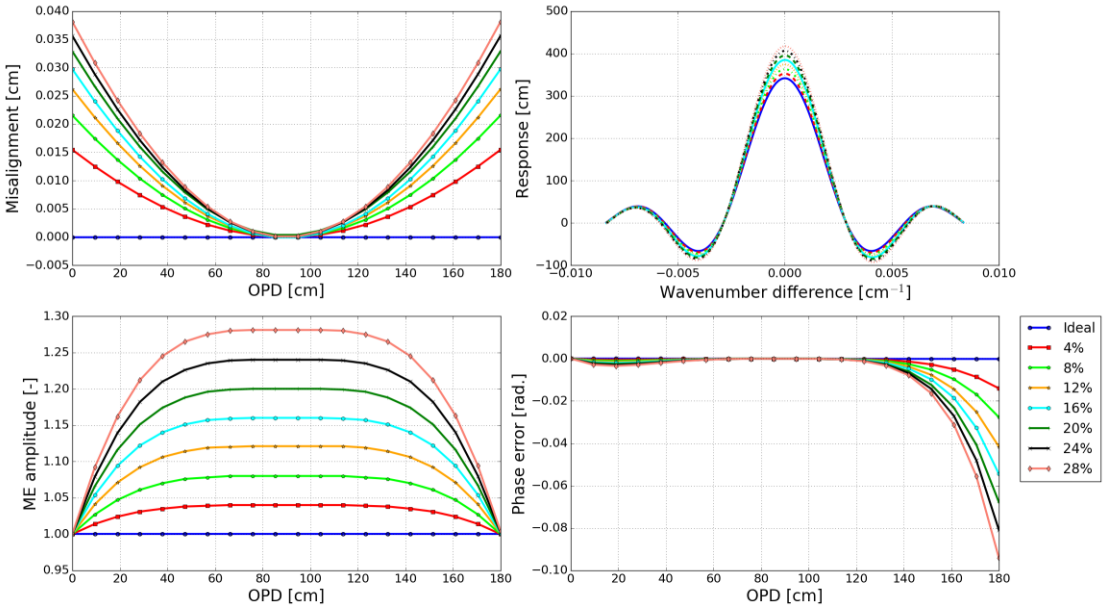

Fig.10. Simulated positive ME deviations along with OPD. Top left demonstrates the misalignment, top right is the resulting ILS, bottom left is the resulting ME amplitude, and bottom right is the resulting PE.

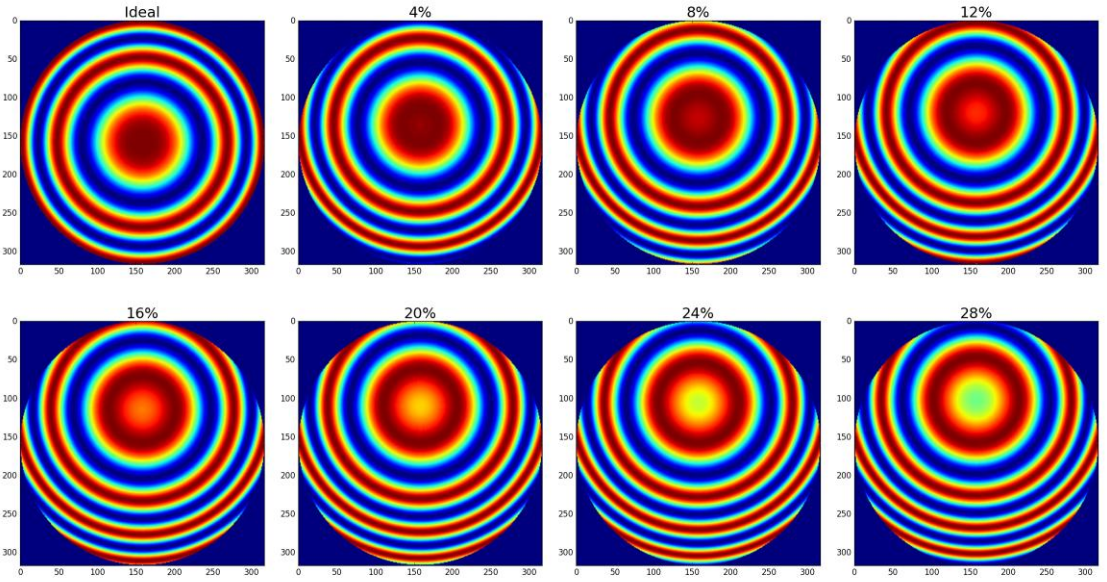

Fig.11. The Haidinger fringes at maximum OPD (the maximum misalignment position) for Fig. 10

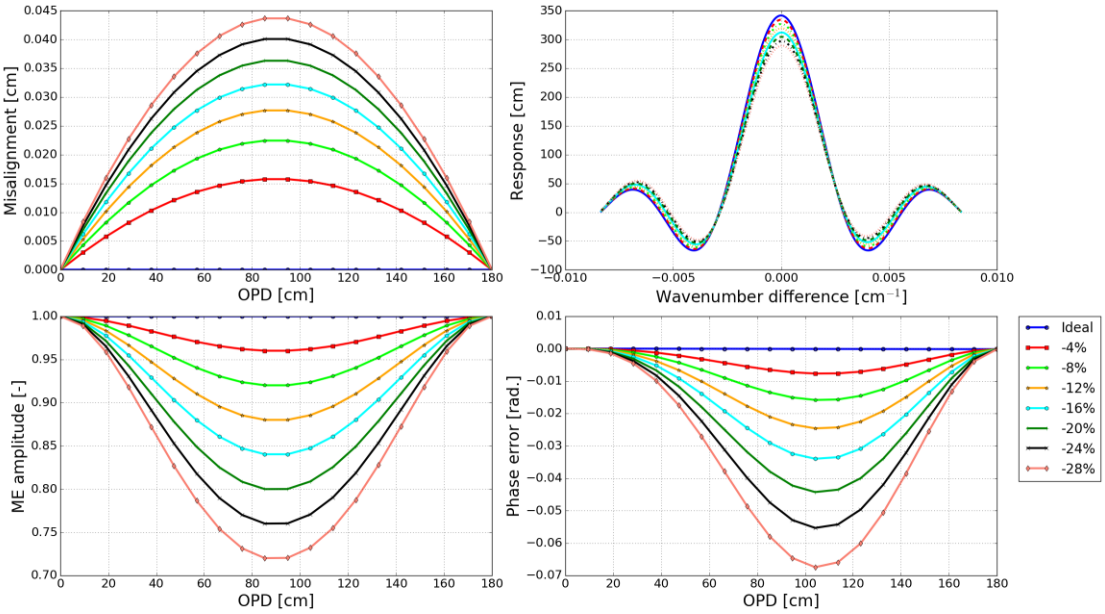

Fig.12. Simulated negative ME deviations along with OPD. Top left demonstrates the misalignment, top right is the resulting ILS, bottom left is the resulting ME amplitude, and bottom right is the resulting PE.

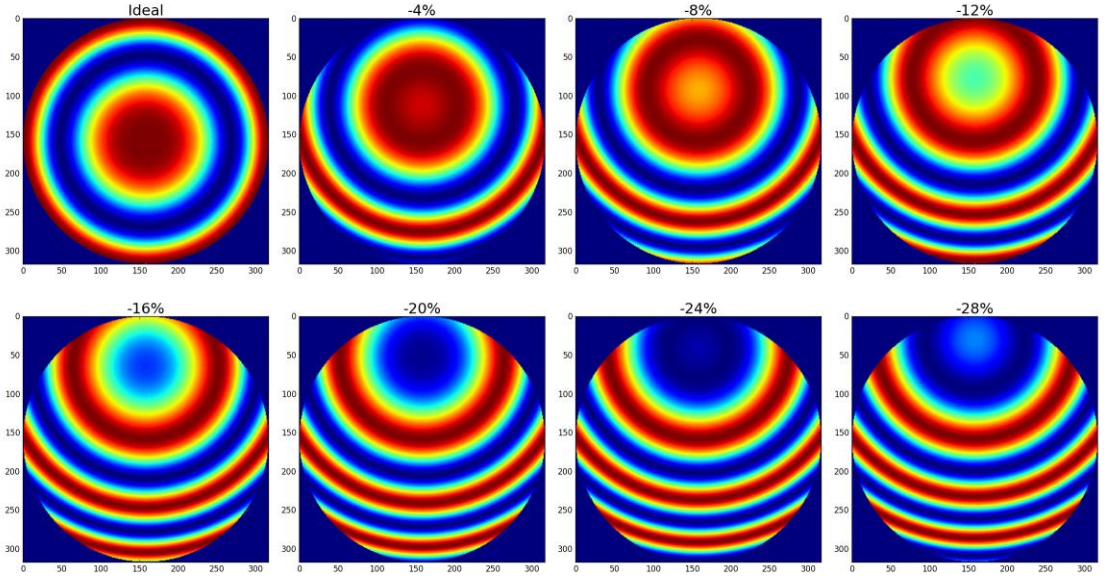

Fig.13. The Haidinger fringes at 1/2 maximum OPD (the maximum misalignment position) for Fig. 12

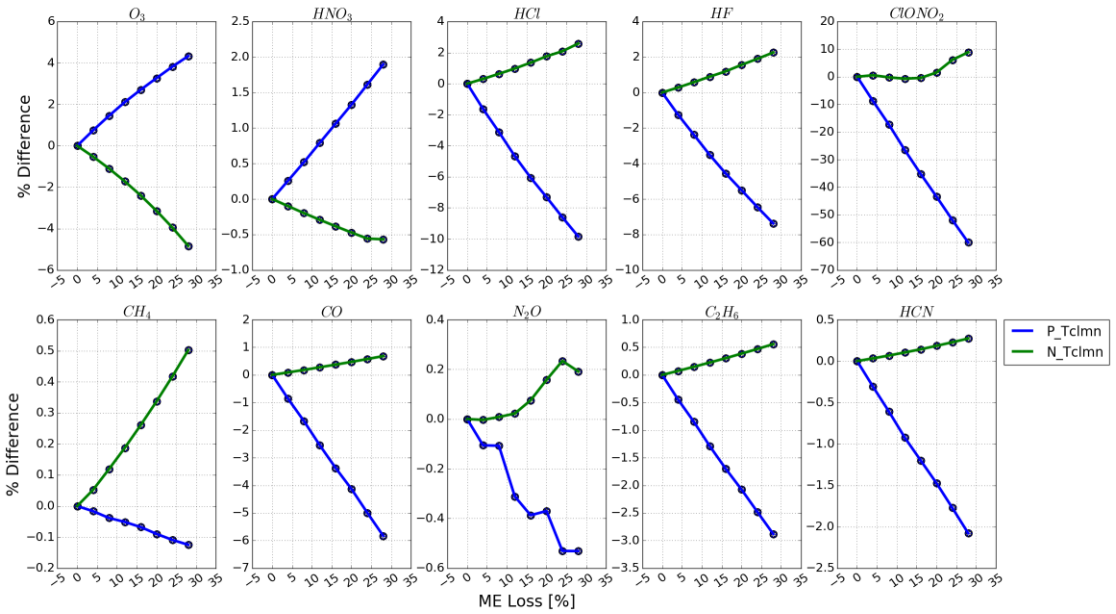

Fig.14. Sensitivity of total column with respect to ME deviation. "P_Tclmn" represents the sensitivity of total column with respect to positive ME deviation and "N_Tclmn" represents the sensitivity of total column with respect to negative ME deviation. The results are deduced from the spectra recorded at Hefei on February 16, 2016.

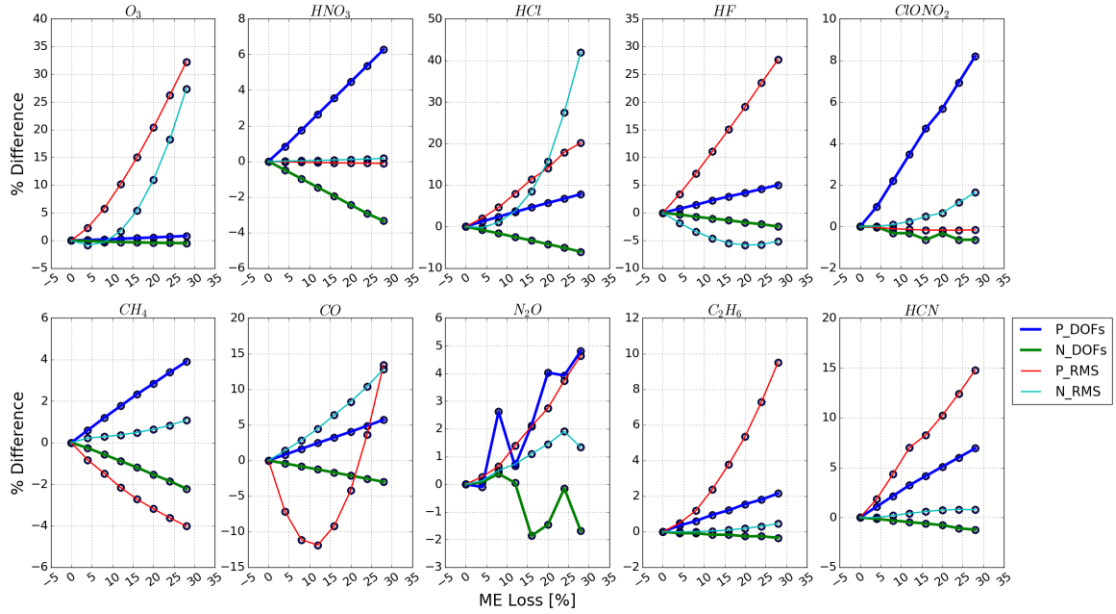

Fig.15. The same as Fig.14 but for DOFs and fitting RMS. The acronyms in the legend are similar to those in Fig.14

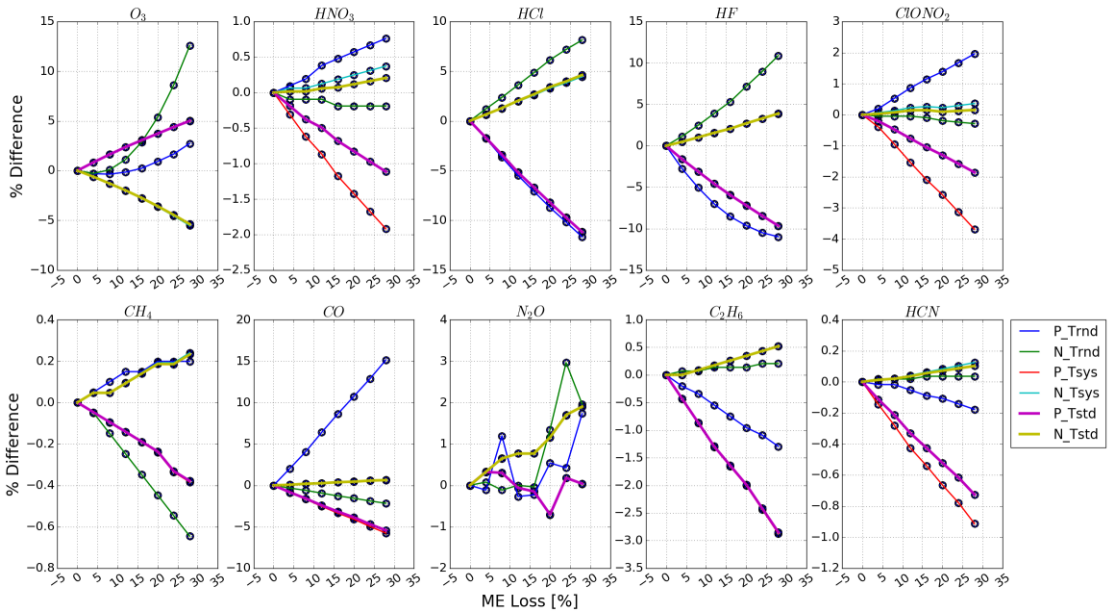

Fig.16. The same as Fig.14 but for total random uncertainty, total systematic uncertainty and total uncertainty. The acronyms in the legend are similar to those in Fig.14. "Trnd", "Tsys" and "Tstd" represent total random uncertainty, total systematic uncertainty and total uncertainty, respectively.

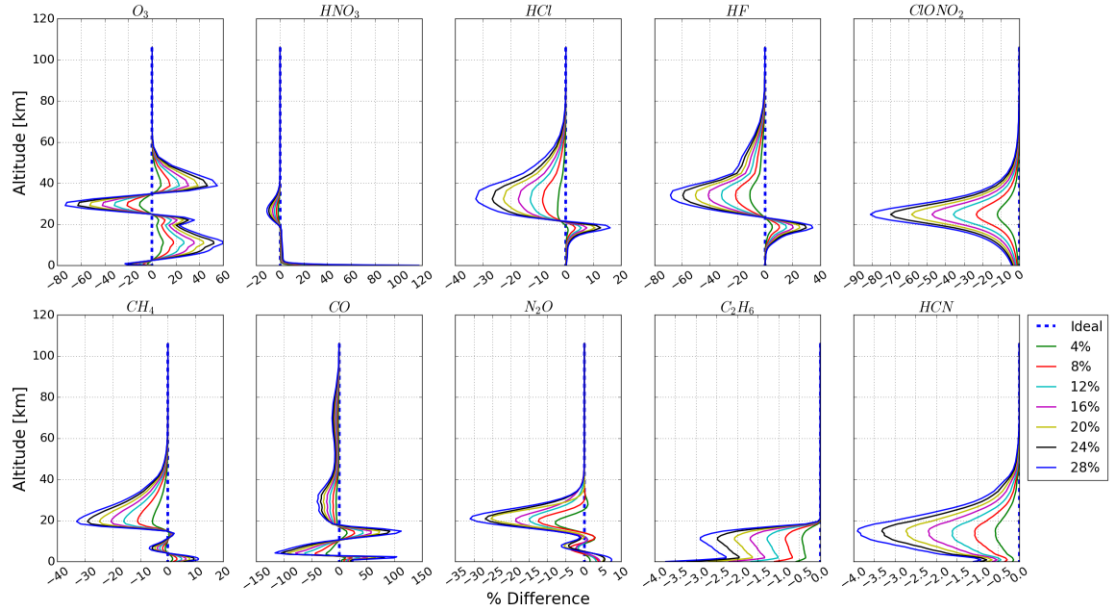

Fig.17. Sensitivity of profile with respect to ME deviation. "4%" represents the ME amplitude deviation is 4%. The nomenclature for other plot labels is straightforward. The results are deduced from the spectra recorded at Hefei on February 16, 2016.

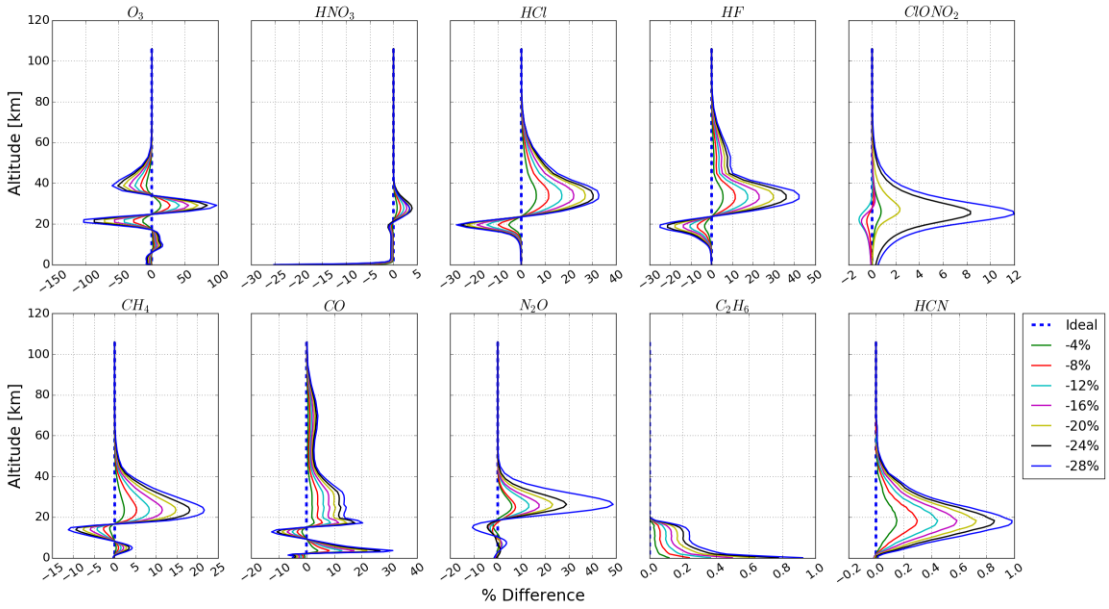

Fig.18. The same as Fig.17 but for negative ME deviation.

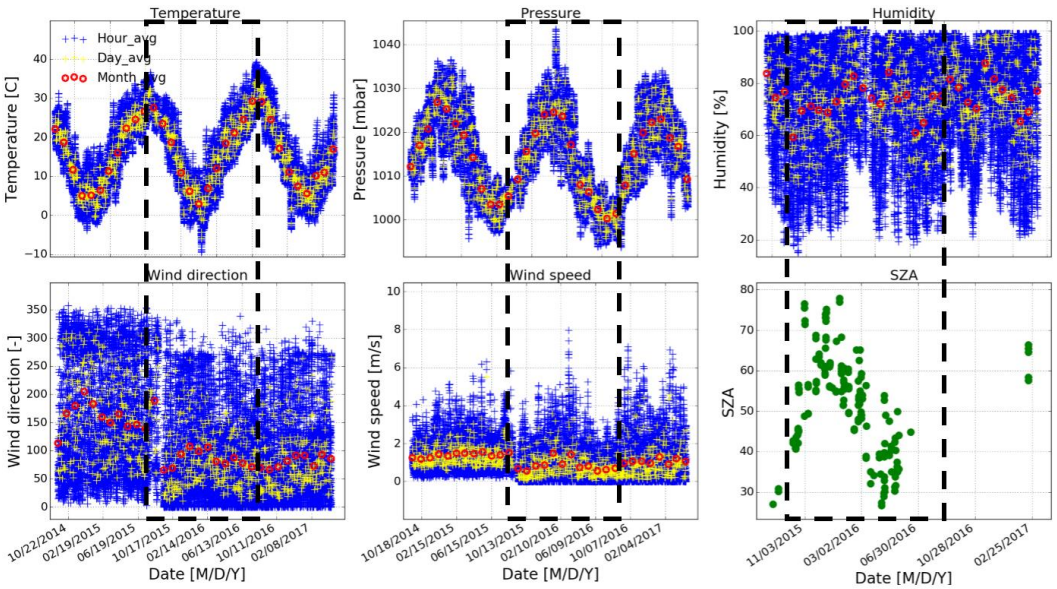

Fig.19. The meteorological data and SZAs record at Hefei. Large span of all these parameters are shown within the period from August 2015 to August 2016 (black dotted square).

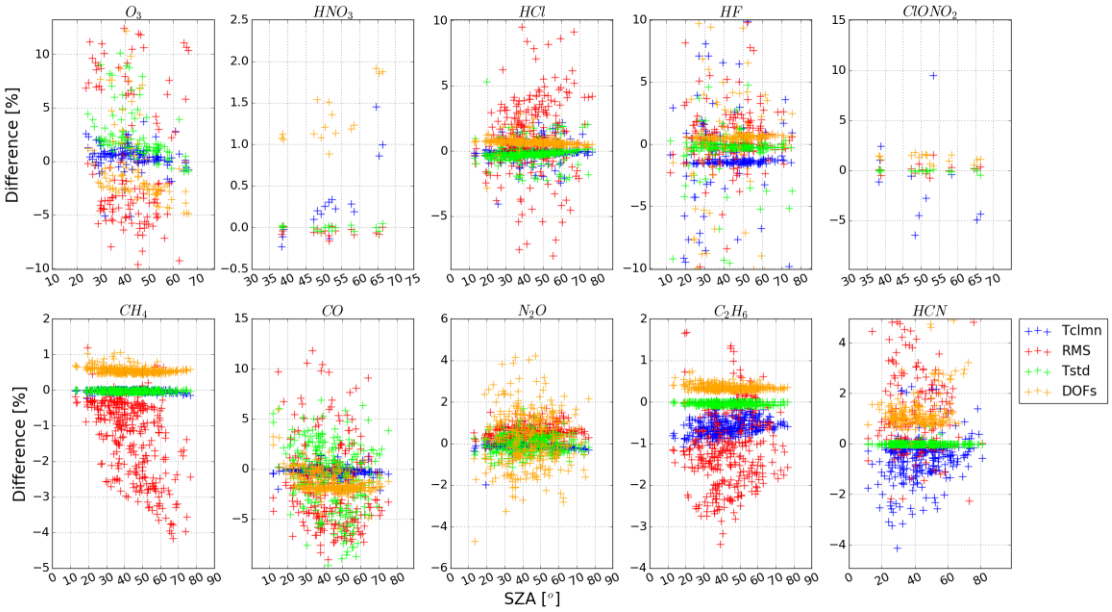

Fig.20. Fractional difference in total column, RMS, total uncertainty, and DOFs as a function of SZA from August 2015 to August 2016 where ILS *j* with a maximum ME deviation of 5% is used.

10 Tables

Table 1. Misalignments simulated in the ALIGN60

Type [a]	Description	Input	Output in maximum
a	No misalignment occurs: interferometer in ideal condition	none	ME amplitude: 1.00 PE: 0.000 rad.
b	Decenter of entrance field stop defining FOV: causes a linear increase in misalignment along OPD	0.33 [mrad] field stop error	ME amplitude: 0.86 PE: -0.056rad.
c	Decenter of path measuring laser: causes a linear increase in phase error along OPD	0.33 [mrad] laser error	ME amplitude:1.00 PE: -0.152rad.
d	Constant shear: causes a constant shear offset of fixed retro-reflector	0.03 [cm]	ME amplitude: 1.00 PE: -0.056 rad.
e	Decreasing linear shear: causes a linear decrease in misalignment along OPD	0.03-0.00017*OPD [cm]	ME amplitude: 1.16 PE: -0.007 rad.
f	Increasing linear shear: causes a linear increase in misalignment along OPD	0.00017*OPD [cm]	ME amplitude: 0.86 PE: -0.056 rad.
g	Cosine bending of scanner bar: causes a cosine decrease in misalignment along OPD	0.03*cos(π*OPD/360) [cm]	ME amplitude: 1.16 PE: -0.013 rad.
h	Sine bending of scanner bar: causes a sine increase in misalignment along OPD	0.03*sin(π*OPD/360) [cm]	ME amplitude: 0.86 PE: -0.056 rad.
i	Cosine & sine bending of scanner bar: causes a chord increase in misalignment before 1/2 maximum OPD and causes a chord decrease in misalignment after 1/2 maximum OPD	0.073*(sin(π*OPD/360)+ cos(π*OPD/360))-0.073 [cm]	ME amplitude: 0.86 PE: -0.029 rad.
j	Constant shear plus cosine & sine bending of scanner bar: causes a chordal decrease in misalignment before 1/2 maximum OPD and causes a chordal increase in misalignment after 1/2 maximum OPD	-0.073*(sin(π*OPD/360)+ cos(π*OPD/360))+0.103 [cm]	ME amplitude: 1.16 PE: - 0.056 rad.

[a] The *b*, *f*, *h*, and *i* are referred to increasing misalignment, the *e*, *g*, and *j* are referred to decreasing misalignment.

Table 2. Summary of the retrieval parameters used for all NDACC gases. All micro windows (MW) are given in cm^{-1}

Gases	O_3	HNO_3	HCl	HF	$ClONO_2$	CH_4	CO	N_2O	C_2H_6	HCN
MW for profile retrievals	1000-1004.5	867.5-870	2727.73-2727.83 2775.7-2775.8 2925.8-2926.0	4109.4-4110.2	779.85-780.45 782.55-782.87	2613.7-2615.4 2835.5-2835.8 2921.0-2921.6	2057.7-2058 2069.56-2069.76 2157.5-2159.15	2441.8-2444.6 2481.2-2482.5	2976-2978 2982.6-2984.5	3268-3268.38 3287-3287.48
Retrieved interfering gases	H_2O, CO_2, C_2H_4, O3668, O3686	H_2O, OCS, NH_3	CH_4, NO_2, O_3, N_2O, HDO	H_2O, HDO, CH_4	O_3, HNO_3, H_2O, CO_2	CO_2, NO_2, H_2O, HDO	O_3, N_2O, CO_2, OCS, H_2O	CO_2, CH_4	H_2O, CH_4, O_3	H_2O, O_3, C_2H_2, CH_4
H_2O treatment	Profile retrieval	Scaling retrieval	Profile retrieval	Profile retrieval	Scaling retrieval	Profile retrieval	Profile retrieval	Profile retrieval	Profile retrieval	Profile retrieval
SNR for de-weighting	None	None	300	None	None	None	500	None	None	None
S_a	Diagonal: 20% No correlation	Diagonal: 50% No correlation	Diagonal: 50% No correlation	Diagonal: 10% No correlation	Diagonal: 100% Exponential correlation HWHM: 8 km	Diagonal: 10% No correlation	Diagonal: 11% ~ 27% No correlation	Diagonal: 10% No correlation	Diagonal: 10% No correlation	Diagonal: 21% ~ 79% No correlation
Error analysis	Systematic error: -Smoothing error -Errors from parameters not retrieved by sfit4 [a]: Background curvature, Optical path difference, Field of view, Solar line strength, Background slope, Solar line shift, Phase, Solar zenith angle, Line temperature broadening, Line pressure broadening, Line intensity									
	Random error: -Interference errors: Retrieval parameters, Interfering species -Measurement error - Errors from parameters not retrieved by sfit4 [a]: Temperature, Zero level									

[a]The input uncertainties of all these items are the same and are included into error analysis if they are not retrieved. Otherwise, the corresponding uncertainties wouldn't be included.

Table 3. Altitude ranges with sensitivity larger than 0.5 for all NDACC gases

Items	O_3	HNO_3	HCl	HF	$ClONO_2$	CH_4	CO	N_2O	C_2H_6	HCN
Altitude ranges (km)	Ground - 44	17 - 28	18 - 42	18-44	20 - 28	Ground - 31	Ground - 27	Ground - 31	Ground - 13.5	4.5-18
Total DOFs	5.2	1.4	1.5	1.3	0.55	3.5	3.8	4.0	1.2	1.1

Table 4. Recommendation for suppressing fractional difference in total column for $ClONO_2$ and other NDACC gases within 10% and 1%, respectively

Items	O_3	HNO_3	HCl	HF	$ClONO_2$	CH_4	CO	N_2O	C_2H_6	HCN
Positive ME	< 6%	<15%	<5%	<5%	<5%	*	<5%	*	< 9%	<13%
Negative ME	< 6%	*	<12%	<12%	*	*	*	*	*	*

*The influence on $ClONO_2$ is less than 10% and on all other NDACC gases are less than 1% even the ILS degrade by an excessively large of 28%, and thus can normally be regarded as negligible.